11-15-2022

Enabling NATO's Collective Defense: Critical Infrastructure Security and Resiliency (NATO COE-DAT Handbook 1)

Carol V. Evans

Chris Anderson

Malcom Baker

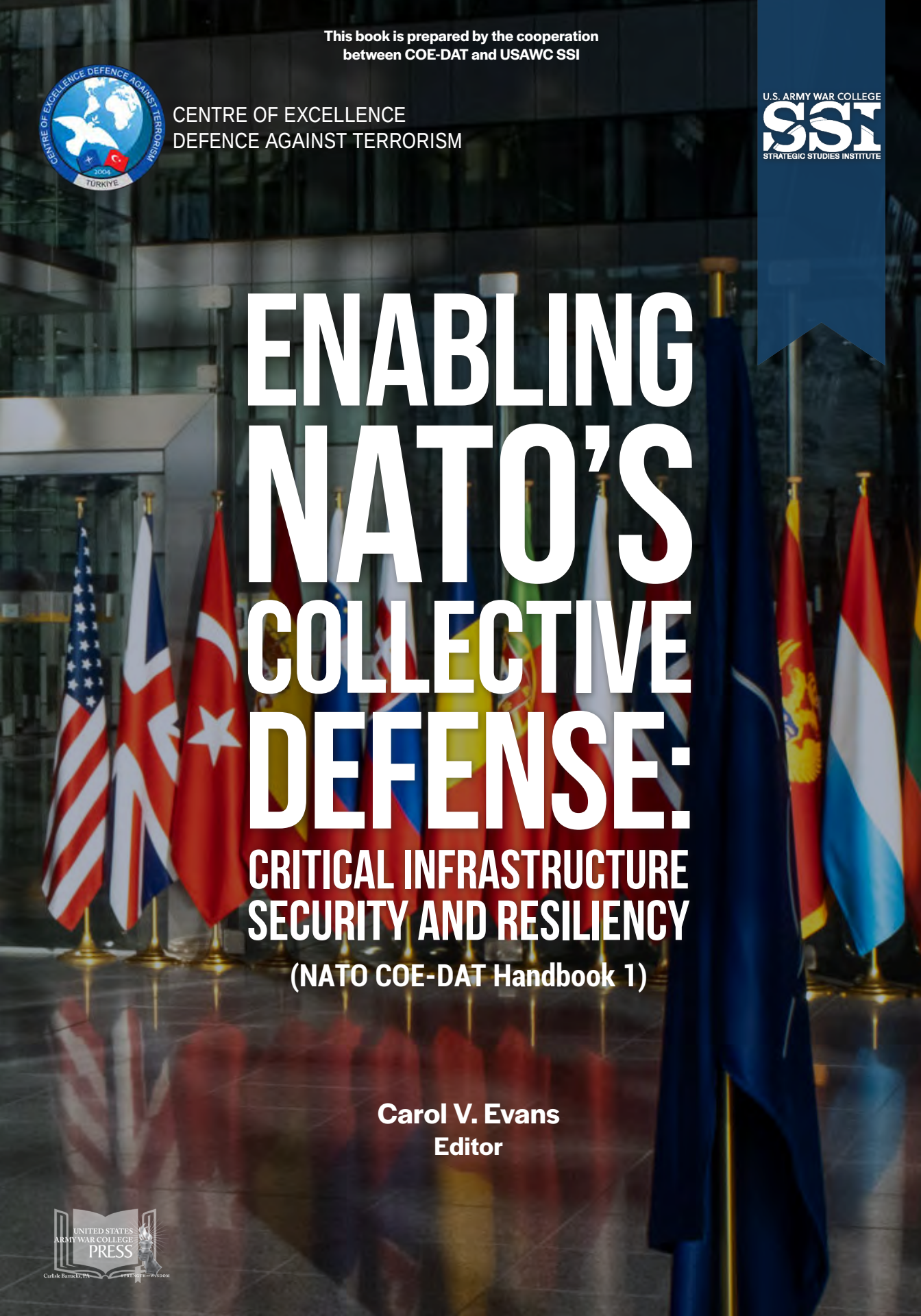
Ronald Bearse

Salih Biçakci

See next page for additional authors

Authors

Carol V. Evans, Chris Anderson, Malcom Baker, Ronald Bearse, Salih Biçakci, Steve Bieber, Sungbaek Cho, Adrian Dwyer, Geoffrey French, David Harell, Alessandro Lazari, Raymond Mey, Theresa Sabonis-Helf, and Duane Verner



This book is prepared by the cooperation between COE-DAT and USAWC SSI

CENTRE OF EXCELLENCE DEFENCE AGAINST TERRORISM

U.S. ARMY WAR COLLEGE SSI STRATEGIC STUDIES INSTITUTE

ENABLING NATO'S COLLECTIVE DEFENSE:

CRITICAL INFRASTRUCTURE SECURITY AND RESILIENCY

(NATO COE-DAT Handbook 1)

Carol V. Evans
Editor

UNITED STATES ARMY WAR COLLEGE PRESS
Carlisle Barracks, PA

STRATEGIC STUDIES INSTITUTE
"The Army's Think Tank"

The Strategic Studies Institute (SSI) is the US Army's institute for geostrategic and national security research and analysis. SSI research and analysis creates and advances knowledge to influence solutions for national security problems facing the Army and the nation.

SSI serves as a valuable source of ideas, criticism, innovative approaches, and independent analyses as well as a venue to expose external audiences to the US Army's contributions to the nation. It acts as a bridge to the broader international community of security scholars and practitioners.

SSI is composed of civilian research professors, uniformed military officers, and a professional support staff, all with extensive credentials and experience. SSI's Strategic Research and Analysis Department focuses on global, transregional, and functional security issues. Its Strategic Engagement Program creates and sustains partnerships with strategic analysts around the world, including the foremost thinkers in the field of security and military strategy. In most years, about half of SSI's publications are written by these external partners.

Research Focus Arenas

Geostrategic net assessment—regional and transregional threat analysis, drivers of adversary conduct, interoperability between partner, allied, IA, commercial, and Joint organizations

Geostrategic forecasting—geopolitics, geoeconomics, technological development, and disruption and innovation

Applied strategic art—warfare and warfighting functions, Joint and multinational campaigning, and spectrum of conflict

Industrial/enterprise management, leadership, and innovation—ethics and the profession, organizational culture and effectiveness, transformational change, talent development and management, and force mobilization and modernization

Enabling NATO's Collective Defense: Critical Infrastructure Security and Resiliency

NATO COE-DAT Handbook 1

Carol V. Evans
Editor

Chris Anderson, Malcolm Baker, Ronald Bearse, Salih Biçakci, Steve Bieber, Sungbaek Cho, Adrian Dwyer, Geoffrey French, David Harell, Alessandro Lazari, Raymond Mey, Theresa Sabonis-Helf, Duane Verner
Contributors

November 2022

UNITED STATES ARMY WAR COLLEGE PRESS
Carlisle Barracks, PA — STRENGTH—WISDOM

Strategic Studies Institute

Comments pertaining to this publication are invited and should be forwarded to: Director, Strategic Studies Institute and US Army War College Press, US Army War College, 47 Ashburn Drive, Carlisle, PA 17013-5244.

******

ISBN 1-58487-840-1

Cover Photo Credits

Front and Back Covers

NATO - Bilateral Meeting with Minister Radmila Shekerinska
Description: Arrival of Minister Radmila Shekerinska
Photo by: NATO
Photo Date: February 13, 2019
Website: https://www.flickr.com/photos/nato/33206419888/

Table of Contents

Preface

The Centre of Excellence for the Defence Against Terrorism (COE-DAT) provides decisionmakers with a comprehensive understanding of terrorism and counterterrorism in order to transform the North Atlantic Treaty Organization (NATO) and nations of interest to meet future security challenges. This transformation is embedded in NATO's three declared core tasks of collective defense, crisis management, and cooperative security. COE-DAT recognizes that counterterrorism is an extremely broad security challenge. COE-DAT also recognizes that military forces alone will not be able to defeat terrorism, nor should military forces be the lead agency in the fight against terrorism. Terrorism evolves from local grievances and as such requires a whole-of-government and whole-of-society approach that includes strategic cooperation and the collective action of nations, civil society, and the international community.

As a strategic-level think tank for the development of Alliance activities to defend against terrorism and sitting outside the formal NATO Command Structure, COE-DAT supports NATO's Long-Term Military Transformation by anticipating and preparing for the ambiguous, complex, and rapidly changing future security environment. COE-DAT is able to interact with universities, think tanks, researchers, international organizations, and global partners to provide critical thought on the inherently sensitive topic of counterterrorism.

This project traces back to the middle of 2019 as part of the after-action review process of COE-DAT's sixth iteration of our Critical Infrastructure Protection Against Terrorist Attacks (CIPATA) course. Recommendations on potential ways to transform the CIPATA course to better serve NATO and partner nations set in motion a series of events that ultimately led to the development of this book. One of the most important events was the adoption of a formal partnership between the US Army War College Strategic Studies Institute (USAWC SSI) and NATO COE-DAT.

USAWC SSI is the US Army's premier strategic-level think tank, tasked to deliver independent, multidisciplinary research and analysis on international security, geostrategy, and other topics for the US Department of Defense and the broader national security and interagency communities. This partnership enables our two organizations to cooperate rapidly and collaborate to provide

research, analysis, and educational support on security issues in support of NATO, its Allies, and partner nations.

We recognized the emerging challenges of the twenty-first-century's security environment are influencing the way nations need to build their concept of critical infrastructure protection. Indeed, the protection of national critical infrastructures has shifted to a risk analysis–based approach focused on developing security and resilience, hence the focus on critical infrastructure security and resilience (CISR).

NATO Allies understand the continuous transformation of securing critical infrastructure because resilience is one of NATO's seven key competencies. Although CISR is primarily an individual Ally's responsibility, protecting key global critical infrastructure is a strategic-level issue and challenge. Considering that global critical infrastructure—such as global transportation, financial and telecommunication networks, and supply chains for energy or medical supplies—is an international security concern, NATO, along with the European Union and the United Nations, needs to be involved.

Although NATO and other nations are currently developing and/or implementing their own strategic and operational approaches to infrastructure protection, a comprehensive approach was needed on how to strengthen national and global, critical infrastructure security and resilience. To answer this need, COE-DAT and USAWC SSI conducted a two-year research effort drawing upon the leading international subject matter experts on critical infrastructure protection from government, academe, and industry. The culmination of this research effort is this handbook, and though there are many books on CIP, there is no joint publication within NATO in terms of critical infrastructure.

The aim of this project, therefore, is to fill this gap by approaching the subject from a military point of view and collecting the best practices in CISR. The book first defines CISR and the many threats to it, presents critical infrastructure–sector case studies, and finally provides tools nations can use to strengthen CISR policies and practices.

Based on the success, energy, and interest garnered during the writing of this first book, a continuation volume two—again in collaboration with the USAWC SSI—began in late 2021. While this first volume sets the stage, the second volume will focus on Article 3, covering the capability and capacity for defense (military) and civil preparedness and resilience. The aim of the second volume is to relate CISR efforts to NATO's seven

baseline requirements. The two volumes individually and collectively are intended to be resources for Allies and partners to develop and improve their respective CISR efforts, but they can also be used as a teaching resource.

COE-DAT, in cooperation with the USAWC SSI, offers this publication to the NATO community, partner nations, other nations of interest, and academia to promulgate "good practices" in the global fight against terrorism. The strategic relationship between the USAWC SSI and NATO COE-DAT led to this book becoming reality.

Daniel W. Stone
Colonel, US Air Force
Deputy Director, COE-DAT
January 2022

Acknowledgments

This book would not be possible without the efforts of the US Army War College Strategic Studies Institute (USAWC SSI) as well as my staff at the Centre of Excellence for the Defence Against Terrorism (COE-DAT).

I want to thank the authors: Chris Anderson, Malcolm Baker, Ronald Bearse, Salih Biçakçi, Steve Bieber, Sungbaek Cho, Adrian Dwyer, Carol V. Evans, Geoffrey French, David Harell, Alessandro Lazari, Theresa Sabonis-Helf, and Duane Verner. Without your valuable expertise and insights into practical solutions to counterterrorism in reference to critical infrastructure security and resilience, this project would not have been possible. I especially want to thank Ronald Bearse, Carol Evans, Michael Lowder, and Geoffrey French who provided the impetus and initial inspiration for the development of this book.

COE-DAT is grateful for our academic collaboration with the USAWC SSI, under the leadership of Dr. Carol Evans, who provided the conceptual framework, academic rigor, and oversight for this endeavor. Many thanks to Dr. Sarah Lohmann for critical reviews and salient inputs that made the book more succinct and well-rounded. I also want to thank Lieutenant Colonel Jeffrey Van Sickle for his tireless copy editorial efforts to bring the book chapters to their on time, polished end states and for his assistance in conducting the series of author workshops. Many thanks to the USAWC Press for their support in publishing this volume.

I am highly indebted to members of my staff: Colonel Daniel W. Stone, Colonel Attila Csurgo, Captain Savaş Aydoğdu, Colonel Pavlin Raynov, Colonel Ioan Pribek, Lieutenant Colonel Uwe Berger, Major Ian McDonald, Major Michael Pasquale, Major James Kadel, and Major Bert Venema. Thank you for your endless patience, tireless work, critical reviews, and passion to complete this book. A special thanks goes to Ms. Selvi Kahraman because without her technical skills we would never have been able to coordinate this project with our team members all around the world.

Oğuzhan Pehlivan, PhD
Colonel (TUR A)
Director, COE-DAT

Executive Summary

In 2014 NATO's Center of Excellence-Defence Against Terrorism (COE-DAT) launched the inaugural course on "Critical Infrastructure Protection Against Terrorist Attacks." As the CIPTA course garnered increased attendance and interest, the core lecturer team felt the need to update the course in critical infrastructure (CI) taking into account the shift from an emphasis on "protection" of CI assets to "security and resiliency." What was lacking in the fields of academe, emergency management, and the industry practitioner community was a handbook that leveraged the collective subject matter expertise of the core lecturer team, a handbook that could serve to educate government leaders, state and private-sector owners and operators of critical infrastructure, academicians, and policymakers in NATO and partner countries. *Enabling NATO's Collective Defense: Critical Infrastructure Security and Resiliency* is the culmination of such an effort, the first major collaborative research project under a Memorandum of Understanding between the US Army War College Strategic Studies Institute (SSI), and NATO COE-DAT.

The research project began in October 2020 with a series of four workshops hosted by SSI. The draft chapters for the book were completed in late January 2022. Little did the research team envision the Russian invasion of Ukraine in February this year. The Russian occupation of the Zaporizhzhya nuclear power plant, successive missile attacks against Ukraine's electric generation and distribution facilities, rail transport, and cyberattacks against almost every sector of the country's critical infrastructure have been on world display. Russian use of its gas supplies as a means of economic warfare against Europe—designed to undermine NATO unity and support for Ukraine—is another timely example of why adversaries, nation-states, and terrorists alike target critical infrastructure. Hence, the need for public-private sector partnerships to secure that infrastructure and build the resiliency to sustain it when attacked. Ukraine also highlights the need for NATO allies to understand where vulnerabilities exist in host nation infrastructure that will undermine collective defense and give more urgency to redressing and mitigating those fissures.

The conceptual framework for this handbook addresses key aspects of which users need to have a baseline knowledge. What is critical infrastructure and why is it important for NATO and an individual nation's security? Threats and attacks to CI may occur from many vectors, including kinetic attacks conducted largely by terrorists to cyberattacks by terrorists, nation-states,

and their proxies, to hybrid threats. Among the many critical infrastructure sectors, there are designated lifelines, ones that are vital due to their importance to the well-being of society, the continuity of government operations, economic impacts, and the deleterious, cascading effects on other CI sectors. There has been a recent shift in the critical infrastructure community of practitioners from an emphasis on merely the "protection" of key vital infrastructure assets to building in "security and resilience" of that infrastructure. What then are the tools that nation-states and owners and operators of CI can employ to achieve these twin goals?

To provide an understanding of these important CI topics, SSI and COE-DAT brought together leading international experts. This multidisciplinary team consisted of industry practitioners, US and European policymakers, members from the intelligence community, research laboratory experts, and academicians. *Enabling NATO's Collective Defense: Critical Infrastructure Security and Resiliency* consists of three major sections. The first section includes four chapters that focus on what we mean by "critical" infrastructure and why and how it has been targeted. There has been an evolution of physical attacks, mainly by terrorists, to sophisticated cyberattacks by adversaries and to more complex hybrid means. Chapter 1, "Understanding Critical Infrastructure," by Ron Bearse sets the stage for the book by answering the following questions: What is critical infrastructure? Why is it important? What is the difference between critical infrastructure protection (CIP) and critical infrastructure security and resilience (CISR)? What is involved in implementing CISR policy in and across the North Atlantic Treaty Organization nations? Bearse suggests that CISR is a quintessential societal task for maintaining national security, economic vitality, and public health and safety in a world filled with increasing levels of risk. For NATO member states, building and enhancing CISR at the national level is necessary to safeguard societies, people, and shared values and also provide the foundation for credible deterrence and collective defense.

Chapter 2, "Physical Threats to Critical Infrastructure," by Malcolm Baker, Ronald Bearse, and Ray Mey details kinetic threats to CI by terrorists with a useful case study regarding the 2013 attack by an al-Qaeda affiliate on the Amenas oil and gas facility in Algeria. They also examine natural and other physical threats to infrastructure, as well as future man-made threats that are of greatest concern to NATO, including chemical, biological, radiological, nuclear, explosive (CBRNE) devices, drones and unmanned aerial vehicles, precision strike weapons, and an electromagnetic pulse attack.

Chapter 3, "Cyber Threats to Critical Infrastructure," by Salih Biçakci examines how risks against critical infrastructure are on the rise in the cyber domain. He writes that while the COVID-19 pandemic has compelled businesses to adopt practices to accommodate a more remote workforce, it has also presented malevolent attackers an unprecedented opportunity to test cybersecurity systems and exploit vulnerabilities. The pandemic has demonstrated the need for the dependable and continuous operation of electricity, natural gas, oil, water and wastewater systems, and telecommunications. His chapter provides an overview of a critical infrastructure's technical layers and systems and its potential organizational vulnerabilities to cyberattacks related to the human workforce and management. He highlights the various categories of threat actors (opportunistic attackers, competitors, insider threats, advanced persistent threats, and hacktivists) and concludes with an overview of recent primary attack types that threat actors employ to exploit vulnerabilities in critical infrastructure.

The fourth and last chapter in section one is "Hybrid Threats to US and NATO Critical Infrastructure" by Carol V. Evans. She provides an analysis of several major hybrid threat vectors to critical infrastructure with the potential to attack, undermine, or compromise US and NATO warfighting, force projection, and sustainment capabilities. The first threat vector is the deliberate cyber infiltration by adversaries of the energy infrastructure that supports US installations and bases. This infiltration enables adversaries to interfere with the US military's ability to deploy and sustain forward combat forces and equipment. A second hybrid threat vector is adversarial targeting of US and NATO logistics, with the potential to degrade US overseas force projection as well as NATO mobility and sustainment within the theater. The third hybrid threat stems from China's strategic penetration, ownership, and control of key defense industrial-base infrastructure and supply chains in Europe via its Belt and Road Initiative and foreign direct-investment activities. This vector provides an opportunity to undermine US and NATO interoperability and political unity. Dr. Evans's chapter concludes by highlighting US and NATO measures to redress and mitigate these threats by investing in CISR through organizational capacity building, development of policy frameworks, and the implementation of host country baseline resilience requirements.

The second section centers on giving readers an appreciation of the critical "lifeline" infrastructure sectors, namely, energy and transport (including civil aviation and mass transit rail), water, and communications. Leading this section is Chapter 5, "European Energy and the Case of Ukraine,"

by Theresa Sabonis-Helf. Written prior to the Russian invasion of Ukraine in February 2022, Dr. Sabonis-Helf posits that potential electricity interruption in the West is becoming both increasingly catastrophic for urbanized areas and more attractive to threat actors seeking disruption. The avenues for disruption are becoming greater as energy systems become larger, "smarter," and more internationally linked. The intertwined relationship between electricity security and cybersecurity calls for an understanding of CISR that recognizes both sets of vulnerabilities. Dr. Sabonis-Helf argues that the case of Ukraine is thus instructive. Ukraine's experience of energy security and cybersecurity reveals significant risks and offers insight into NATO's efforts to enhance civil preparedness and collective CISR among Allies and partners. It also illustrates the complexities that Ukraine and Europe are facing today and will face in the future.

Chapter 6, "Civil Aviation," by David Harell analyzes the aviation infrastructure sector and the threats it faces, including primary aircraft bombings and ground attacks on airports by terrorists. To understand the civil aviation sector, he writes it is important to know why the aviation industry is so critical, what makes it so volatile, and why it is such an attractive target for terrorists. He provides several key reasons for the industry's vulnerability: its rigidity, its predictability, and its difficulty in keeping up with evolving terrorist threats. Harell uses multiple case studies—which span the 20 years after the 9/11 attacks—to illustrate these vulnerabilities. He also examines the aviation security responses to these terrorist attacks and identifies important lessons to be learned. He concludes with recommendations and best practices that can assist in reducing the vulnerabilities across international civil aviation.

Chapter 7, "Mass Transit Railway Operations," by Adrian Dwyer explains the inherent vulnerability of open transport networks, such as railway operations, to terrorist action. He showcases those methods of attack that have often been used, drawing on case study data from Great Britain, continental Europe, the United States, Japan, and India. From the perspective of NATO, the targeting of rail networks across its member states can disrupt military logistics, the civilian supply chain, and economic prosperity more generally. Dwyer maintains that strategic risk assessment is an important means to manage diverse terrorist threats and inherent vulnerabilities of mass rail transit.

Chapter 8, "Water Sector Resilience and the Metropolitan Washington Case," by Steve Bieber provides an eye-opening analysis of how fragile the supply of this vital resource can be and how other sectors of critical infrastructure are highly dependent on water. Bieber identifies the risks

and threats to the water sector; outlines key steps in resilience planning; illustrates challenges and solutions to security and resilience initiatives using a case study from Washington, DC; and offers recommendations for developing water-sector security and resilience. He explains that the security and resilience of the water sector is a key enabler of a nation's civil preparedness, with military implications as well. Terrorist threats to water delivery or contamination of water sources can impact a nation's ability to move and sustain its military forces and project military power when required. From the perspective of the North Atlantic Treaty Organization, threats to the water sector in one member state could have ripple effects that limit or diminish NATO's military mobility and force projection in support of its essential core tasks.

Chris Anderson's Chapter 9, "Communications Resilience," completes the handbook's second section on the lifeline CI sectors. Communications form the critical backbone of the modern world, and resilient and trustworthy communications are fundamental to national security and emergency preparedness. Communications play many critical roles for NATO, he writes, including: command and control, military operations, distribution of intelligence and warning signals, crisis management and coordination, and citizen preparedness. Anderson provides an in-depth overview of the communications sector and explains the ways in which the integrity, availability, or confidentiality of communications systems may be degraded or compromised. He shows the risks to the communications sector using recent natural, man-made, cyber, and kinetic incidents that have impacted communications systems and related infrastructure. He provides important recommendations for improving communications resilience against terrorist attacks and other threats.

The third section of *Enabling NATO's Collective Defense: Critical Infrastructure Security and Resiliency* provides readers and users of the handbook with the tools necessary to deter and mitigate attacks against critical infrastructure as well as the means to build long-term security and resiliency within host nation infrastructure, thereby enabling NATO's collective defense. There are six chapters in this section. It begins with Chapter 10 by Ron Bearse and Alessandro Lazari, who employ their respective policy-making purviews to collaborate on "Comparing Policy Frameworks: CISR in the United States and the European Union." The US and EU CISR policies and practices are the most advanced frameworks in the world, and many countries have emulated the US and EU models. Their chapter describes the key underpinnings and characteristics of each respective policy framework, the reasons why these frameworks came into being, and how they were adapted over time. The intent of this chapter is to help Allies and partners better

understand these two CISR policy frameworks so they can apply the key principles and tenets to enhance the CISR posture in their respective countries.

In Chapter 11, "Information and Intelligence Sharing," Chris Anderson and Raymond Mey discuss the important role of information and intelligence sharing between governments and state or private sector owner operators of critical infrastructure. These activities are essential to the success of any CISR effort across the North Atlantic Treaty Organization. They explain that key infrastructure stakeholders need to share information to understand comprehensive infrastructure risk so they can then determine the most efficient and effective means to mitigate these dangers. This process involves building trust, shared and practiced communications methods, and structured, multidimensional information sharing. Anderson and Mey provide some best practices of public-private information-sharing programs from the United States. These practices include the Department of Homeland Security Critical Infrastructure Partnership Advisory Council (DHS CIPAC), Protected Critical Infrastructure Information (PCII) program, and Cyber Information Sharing and Collaboration Program, and the Federal Bureau of Investigation (FBI) Joint Terrorism Task Force (JTTF), the Domestic Security Alliance Council (DSAC), and InfraGard. One information-sharing program common in the United States and Europe is an Information Sharing and Analysis Center (ISAC), which is a critical infrastructure sector-specific organization to share information about threats and vulnerabilities.

Chapter 12, "Critical Infrastructure Interdependency Modeling and Analysis: Enhancing Resilience Management Strategies," by Duane Verner provides a tour-de-force overview of the need for NATO member states and partner nations to understand infrastructure interdependencies since they operate in concert with each other. Catastrophic events can cascade across these interconnected systems and hamper the ability of critical infrastructure operators to remain operational. Modeling and analysis of these interdependencies are key components to an effective risk-management strategy and to determining where resources are needed to build resiliency. Verner summarizes general approaches to model and assess critical infrastructure, and he proposes a flexible CISR framework to inform the development of resilience management strategies. NATO Allies and partners can use this framework to reduce the risks posed to critical infrastructure and to foster greater resilience through cross-sector collaboration.

How NATO can best manage and assess security risk in a constantly changing environment is the starting point for Geoffrey French's Chapter 13

on "Security Risk Assessment and Management." As he points out, organizations and communities need formal processes to determine, prioritize, and address risks. The promise of risk management is that with sufficient uniformity and consistency, government or private-sector leaders can make better decisions through the ability to aggregate risks at different levels. His chapter explores in depth the concepts of risk assessment and risk management and reviews a set of selected risk-management frameworks—from the International Organization for Standardization (ISO), NATO, US Government Accountability Office (GAO), and the US National Infrastructure Protection Plan (NIPP)—that have been designed or adapted for security risk management. French then demonstrates how a national-level governmental risk program can encourage and guide risk-management practices as well as coordinate the constellation of public- and private-sector organizations involved in critical infrastructure operations to foster a mutually supportive environment for CISR.

While Biçakci's earlier chapter describes the various cyber threats to CI, Sungbaek Cho provides users of this handbook with some important cybersecurity tools to mitigate those very threats. In Chapter 14, "Enhancing Cybersecurity of Industrial Control Systems," Cho offers a brief overview of the characteristics of industrial control systems (ICSs) and why they are subject to cyberattacks in terms of the vulnerability of the components as well as the prevalent practice in modern critical infrastructure to operate ICS in more open interconnections with business networks. He highlights these vulnerabilities with some major cyber event case studies, including: Stuxnet (2010), BlackEnergy (2011), Ukraine Blackout (2015), RWE's Nuclear Power Plant in Germany (2016), TRITON (2017), and the Colonial Pipeline (2021). The chapter offers best practices and tools for critical infrastructure stakeholders, owners, and operators to protect their systems and enhance security and resilience against cyberattacks. Cho recommends the utilization of risk management methodologies, basic hygiene practices, and essential cybersecurity measures. Although NATO is taking steps to improve its collective ability to defend against and respond to cyberattacks against Allied critical infrastructure, individual member states form the first line of defense. Cho suggests national governments should establish mandatory cybersecurity requirements for critical infrastructure—ensuring owners and operators comply with these requirements—and provide security advice as needed. He also advocates for establishing an institutional cooperation mechanism (such as a public-private critical infrastructure security council and a joint cyber response team) so the CI stakeholders' unique capabilities can be integrated at the national level.

Malcolm Baker provides the final chapter in the "tools to build CISR" section and fittingly it focuses on "Crisis Management and Response." Crisis management is an essential component of the Alliance's Strengthened Resilience Commitment announced in June 2021 as part of the NATO 2030 initiative. Baker, however, asks whether the Alliance's current philosophy of crisis management is keeping up with mainstream developments in contemporary crisis management and thought leadership. Further, within the construct of CISR efforts, is NATO's crisis management approach still fit for its purpose—or could it be improved? For NATO, understanding crisis management exclusively in terms of armed conflict and other hostilities may no longer be appropriate or optimal, he suggests (especially in light of the various physical, cyber, and hybrid threats outlined earlier in chapters 2–4). Baker recommends that effective CISR measures can be improved by developing and implementing robust crisis management structures and processes. The key elements of effective crisis management are early warning, an effective strategy, good communication, leadership, and swift decision making. Baker offers a proven crisis-management framework, based on the British Standards Institution, using a staged approach of: anticipate and assess, prepare, response, and recovery. Finally, he reviews new developments in resilience and crisis management and offers suggestions for how NATO could better align its activities to support NATO 2030.

Carol V. Evans

Dr. Carol V. Evans
Editor

Director, Strategic Studies Institute
and US Army War College Press
US Army War College
November 2022

— 1 —

Understanding Critical Infrastructure

Ronald Bearse

This chapter sets the stage for this book by answering the following questions: What is critical infrastructure? Why is it important? What is the difference between critical infrastructure protection (CIP) and critical infrastructure security and resilience (CISR)? What are some of the key terms defined in national CISR policy? What are the core areas of activity or work streams involved in implementing CISR policy in and across the North Atlantic Treaty Organization nations? The answers to these specific questions provide the contextual basis for understanding why CISR is a quintessential societal task for maintaining national security, economic vitality, and public health and safety in a world filled with increasing levels of risk. For NATO member states, building and enhancing CISR at the national level is necessary to safeguard societies, people, and shared values and also provide the foundation for credible deterrence and defense and the Alliance's ability to fulfill its core tasks of collective defense, crisis management, and cooperative security.[1]

What Is Critical Infrastructure?

Although there is no standard or universal definition of critical infrastructure, many Western nations have essentially defined the term as the physical and cyber systems and assets that are so vital to the country

1. "Strengthened Resilience Commitment," NATO (website), June 14, 2021, https://www.nato.int/cps/en/natohq/official_texts_185340.htm.

that their incapacity or destruction would have a debilitating impact on the nation's physical or economic security or public health and safety.[2] Certain socioeconomic activities are vital to the day-to-day economic functioning and security of nations. While there is less consensus on which sectors qualify as critical infrastructure, most countries with an established national CIP or CISR policy identify some or all of the sectors listed in figure 1-1 as critical infrastructure.[3] Most people know critical infrastructure on a daily basis as the power used in their homes, the water they drink, the transportation they rely on for freedom of movement, and the systems they use to communicate with family, friends, and coworkers.[4]

Figure 1-1. List of commonly identified critical infrastructure sectors
(Diagram by TSAT)

To illustrate the relationships between critical infrastructure sectors, figure 1-2 uses blue and red boxes to identify and distinguish various sectors.[5] The red boxes—transportation, water, energy, and communications—are known as lifeline sectors. Given their unique nature, there are four main characteristics that distinguish lifeline sectors from other sectors of critical infrastructure.[6] First, lifeline sectors provide necessary services and goods

2. Cybersecurity and Infrastructure Security Agency (CISA), *A Guide to Critical Infrastructure Security and Resilience* (Washington, DC: CISA, 2019), 4, https://www.cisa.gov/sites/default/files/publications/Guide-Critical-Infrastructure-Security-Resilience-110819-508v2.pdf.

3. "Critical Infrastructure Sectors," TSAT (website), accessed September 27, 2021, https://tsat.net/market/critical-infrastructure/critical-infrastructure-sectors/.

4. CISA, *Guide to Critical Infrastructure*, 4.

5. Department of Homeland Security (DHS), *2015 Energy Sector-Specific Plan* (Washington, DC: DHS, 2015), 19, https://www.cisa.gov/sites/default/files/publications/nipp-ssp-energy-2015-508.pdf.

6. National Infrastructure Advisory Council (NIAC), *Strengthening Regional Resilience* (Washington, DC: NIAC, 2013), 18, https://www.cisa.gov/sites/default/files/publications/niac-regional-resilience-final-report-11-21-13-508.pdf.

that support most homes, businesses, communities, and levels of government. Second, they deliver services that are commonplace in everyday life, but disruption of the service has the potential to develop life-threatening situations. Third, lifeline sectors involve complex physical and electronic networks that are interconnected within and across multiple sectors. Finally, a disruption of one lifeline sector has the potential to affect or disrupt other sectors, creating cascading or escalating failures. The blue boxes along the outer ring illustrate the other critical infrastructure sectors that typically depend on the lifeline sectors for continuous operations.

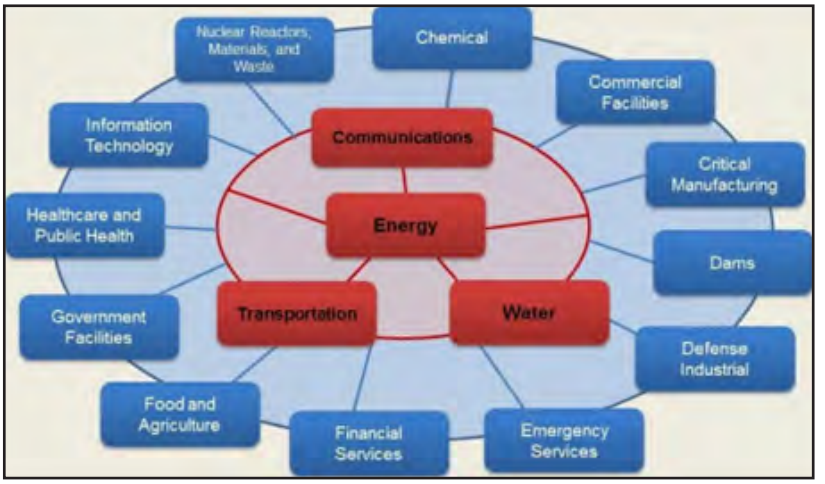


Figure 1-2. Relationships between critical infrastructure sectors
(Diagram by US Department of Homeland Security)

The previous section described how most nations define critical infrastructure, but how does NATO define it? Within NATO, Allied Command Operations (ACO) uses several definitions to identify and understand the types of infrastructure available within a given area of responsibility. ACO defines critical infrastructure as "a nation's infrastructure assets, facilities, systems, networks, and processes that support the military, economic, political and/or social life on which a nation and/or NATO depends."[7] ACO also describes three subcategories of critical infrastructure based on the respective level of impact to national services and/or NATO operations, as listed below.[8]

7. Allied Command Operations (ACO), *Infrastructure Assessment, ACO Directive 084-002* (Mons, Belgium: ACO, 2019), 4.

8. *ACO Directive 084-002*, 4.

- Critical national infrastructure: Assets, facilities, systems, and networks identified by the territorial host nation that are integral to the continued delivery and integrity of the essential services upon which the nation relies, the destruction or compromise of which would lead to severe military, economic, political, or social consequences to the nation.
- Mission vital infrastructure: Assets, facilities, systems, and networks within the joint operations area which NATO/troop contributing nation forces rely on for fielded capability, the destruction or disruption of which singularly creates a decisive disadvantage to the NATO mission.
- Key infrastructure: Assets, facilities, systems, and networks within the joint operations area which host nation or NATO/troop contributing nation forces rely on for fielded capability, the destruction or disruption of which, either singularly or collectively, creates a significant disadvantage to the host nation or NATO mission.

An important question NATO must consider is to what extent its overall mission readiness depends on the assured availability of critical infrastructure, most of which is owned by private sector companies in its various member states. Today, and for some time now, the answer to this question is that NATO depends considerably on this assured availability of critical infrastructure. During large operations or exercises, for example, an estimated 90 percent of military transport relies on civilian ships, railways, and aircraft. See chapter 4 for a discussion on how US and NATO reliance on assured access to the global shipping infrastructure poses threats to Allied force projection and military mobility.

The chapter thus far has discussed how Western nations define critical infrastructure and why they do so. Simply, without the goods and services these critical infrastructure sectors provide, the consequences can be catastrophic for a nation's public health and safety, environment, national security, and/or economy. Furthermore, since many such systems and networks operate across borders, any threat or attack against them could have national, regional, and even global implications. What are some of the other reasons why critical infrastructure is important?

Why Is Critical Infrastructure Important?

Today, citizens in many countries demand or expect critical infrastructure systems and the functions they fulfill to be available 24 hours a day. Adversaries, however, are penetrating and disrupting various parts of critical infrastructure with little or no repercussion. One such example occurred in May 2021 when a small group of hackers launched a ransomware attack on Colonial Pipeline, the largest pipeline network in the United States for delivery of refined petroleum products. See chapter 14 for more detail on the Colonial Pipeline case. Colonial shut down its main lines for five days, disrupting half the fuel supply for the eastern part of the country. Worried drivers drained supplies in gas stations in the southeastern states, airlines rerouted flights to airports with available fuel, traders had to deal with unexpected price volatility, and companies scrambled to locate new sources of fuel.[9]

The Colonial Pipeline attack reveals the importance of building resilience. Events like this one are extremely difficult, if not impossible, to predict, but much can be done to prepare for them. Organizations not only need to improve the security of their systems, but also their ability to respond to an incident and spring back quickly after a disruption. These capabilities can identify in advance the actions to take in response to a large disruptive event.[10] Organizations must know what to do, develop the capabilities to do it, and then rehearse their crisis response actions—all in advance of an incident.[11]

Evidence also suggests that the cyberattack against Colonial Pipeline was not particularly sophisticated, yet it managed to paralyze a significant part of the fuel supply of the world's largest economy. The Colonial Pipeline attack and many similar attacks before and since are made possible by the ongoing automation of traditional manufacturing and industrial practices, making these domains more vulnerable while not necessarily posing a direct threat to them. The fourth industrial revolution's emphasis on modern smart technology is driving societies toward more intelligent and smarter operational networks in domains like energy, water, traffic management, air traffic control, and defense systems, to name a few.[12]

9. Rich Isenberg et al., "Building Cyber Resilience in National Critical Infrastructure," McKinsey & Company (website), June 30, 2021, https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/building-cyber-resilience-innational-critical-infrastructure.

10. Isenberg et al., "Building Cyber Resilience."

11. Isenberg et al., "Building Cyber Resilience."

12. Isenberg et al., "Building Cyber Resilience."

Advances in automated technology—such as sensors, the Internet of Things, and cybersecurity protections—can help those responsible for securing critical infrastructure understand potential threats, run diagnostics, predict potential changes in their systems, and strengthen the security and resilience of these critical infrastructure systems more efficiently. Together, these characteristics of the current threat environment represent one of the most serious national security concerns since the development of the atomic bomb, and they demonstrate why writing a handbook such as this one is necessary to enhance collective CISR posture.

Today, countries are trying to understand the threats they face—some of which are seen and others unseen—and what they might be missing. Those responsible for CISR constantly engage in a process to figure out which threats should be at the top of the list, how others should be ranked, and, given that ranking, what resources they should allocate against those threats. What models and tools are useful to better understand the threats and consequences of attacks on critical infrastructure? Which practices or solutions should countries implement to prevent, deter, mitigate the consequences of, respond to, and recover from events that can disrupt or destroy critical infrastructure? What are the intelligence gaps, and how should stakeholders fill those gaps?[13]

The reality is that any number of factors can cause disruption or destruction to critical infrastructure, including poor design, operator error, natural causes (such as earthquakes, lightning strikes, space weather, and climate change), or intentional human actions like theft, arson, terrorism, or criminal attack. Other factors increasing the risk to critical infrastructure include:

- Poor infrastructure and lack of adequate security controls
- Increased use of information and telecommunication technologies to support, monitor, and control critical infrastructure functionalities
- Growth of the world's population and its transition from rural to urban areas, which stresses the utilization of older infrastructure to its limits

13. James B. Comey, "Protecting Critical Infrastructure and the Importance of Partnerships" (address, Miami, FL, July 7, 2014), https://www.fbi.gov/news/speeches/protecting-critical-infrastructure-and-the-importance-of-partnerships.

- Growing dependencies and interdependencies across critical infrastructure sectors (see chapter 12 for more in-depth analysis on these relationships)
- Adversaries' enhanced understanding that a successful attack to critical infrastructure systems can create havoc
- Nations' dependence on critical infrastructure that is partially or completely located outside their jurisdictional authority and over which they have little or no control

Several chapters in this book delve more deeply into the subjects of vulnerability, threat, and risk. See chapters 2–4 for an overview of threats and threat actors. Nevertheless, it is instructive at this point to share the poignant observation made by information security expert Ben Rothke, who asserted in 2013: "The biggest threat to critical infrastructure is the result of decades of insecurity; combined with an inadequate response to current known threats and vulnerabilities."[14] Rothke's observation continues to haunt governments and companies alike because it underscores the urgent need for all critical infrastructure stakeholders to foster the type of cooperation, coordination, collaboration, communication, and concentration required to harness the expertise necessary to strengthen collective CISR posture demonstrably.

What Is the Difference between CIP and CISR?

Humankind has been protecting critical infrastructure since the invention of the wheel; however, over the last 20 years, most national critical infrastructure policies and strategies in the West have evolved from focusing solely on protection of critical infrastructure to making it more secure and resilient. This shift is primarily due to the impossible task stakeholders face in trying to protect all critical infrastructure systems from the growing number of risk factors they face.

Under the CISR construct, the terms *security* and *resilience* certainly support the idea of protection, but they have specific meanings. *Security* means reducing the likelihood of successful attacks against critical infrastructure or the effects of natural or man-made disasters, through the application of physical means or defensive cybersecurity measures. *Resilience* is the ability of critical infrastructure to resist, absorb, recover from, or successfully adapt

14. Helena Brito, "What Is the Biggest Threat to Critical Infrastructure?," *FireEye* (blog), July 10, 2013, https://www.fireeye.com/blog/threat-research/2013/07/biggest-threat-critical-infrastructure.html.

to changing conditions. Resilient infrastructure is robust, agile, adaptable, and able to withstand and recover rapidly from disruptions, deliberate attacks, accidents, or naturally occurring threats or incidents.

Given the increasing natural and man-made threats and vulnerabilities modern societies face—which have exposed the limits of traditional risk assessment and risk mitigation efforts—the concept of CISR seems particularly useful to inform policies that mitigate the consequences of such events and speak to the vital need for nations to develop and implement a comprehensive risk management strategy. See chapter 13 for analysis and recommendations regarding risk assessment and management strategies.

Transitioning from the goal of protecting to that of improving security and resilience requires a change in focus of education and training to ensure that core CISR work streams are completed and well-managed.

Key Work Streams in CISR Planning and Operations

While the process of establishing, implementing, and sustaining a demonstrably effective national CISR policy is difficult in any country, there are three essential tasks: assessing risk, improving security, and enhancing resilience. The process of accomplishing these three tasks is extraordinarily complex and a continuing challenge because it requires numerous streams of work be performed by all relevant stakeholders: multiple government agencies, critical infrastructure owners and operators, academicians, subject-matter experts, international organizations, and technology vendors. Some of the major work streams include:

- Establishing clear roles and responsibilities for all stakeholders
- Identifying and determining the criticality of national infrastructure
- Mapping critical infrastructure dependencies and interdependencies
- Determining critical infrastructure vulnerabilities
- Using applicable risk assessment, analysis, and management approaches
- Establishing crisis management capabilities

- Establishing public-private partnerships between government and private sector owners and operators of critical infrastructure
- Establishing and implementing collaboration and information sharing mechanisms between government and critical infrastructure owners and operators
- Developing and exercising continuity of operations and information technology disaster recovery plans
- Providing physical and cybersecurity and resilience measures
- Ensuring the integrity, security, and continuity of critical infrastructure supply chains
- Fostering the local, regional, national, and international cooperation, collaboration, coordination, communication, and concentration required to produce demonstrably effective results
- Expanding opportunities to develop and deliver CISR education and training
- Implementing a robust test, training, and exercise program to determine the extent to which current CISR policy, legislation, plans, procedures, systems, and research and development efforts are meeting, falling below, or exceeding prescribed requirements, established standards, and increasingly stringent stakeholder expectations

To implement a national CISR policy successfully requires great leaders and top-notch managers. While the work streams identified above define much of what needs to be done, the extent to which a nation effectively develops and implements "the what" is a function of how well the people responsible for leading and managing these work streams foster the collaboration, cooperation, coordination, communication, and concentration (5Cs) that are indispensable to building and sustaining a viable risk-based, CISR posture. This fact cannot be emphasized enough. For over 30 years, CIP and CISR capabilities have waxed and waned depending on how well leaders execute these 5Cs.

Looking Back and Looking Ahead

While it is important to look ahead, NATO member states can also learn from the past. One instructive example from the Roman Empire is the link between deteriorating conditions of the vast Roman aqueduct system—certainly an early example of critical infrastructure—and the eventual fall of the empire. In the third century BC, the Romans constructed aqueducts throughout the empire to bring water from outside sources into cities and towns to supply public baths, latrines, fountains, and private households, and to support mining operations, milling, farms, and gardens. However, as the administration of the day gradually neglected the aqueducts and did not implement sufficient security and resilience measures against an array of threats and vulnerabilities (to include general wear and tear, intentional rerouting by local farmers, or deliberate destruction or hindering by enemies) many aqueducts ceased to function.[15]

In the twentieth century, threats to critical infrastructure were almost exclusively tangible, physical threats which could be countered with tangible, physical defenses. Those kinds of tangible, physical threats continue today—as do natural disasters, such as hurricanes, floods, and wildfires—and they can cause serious harm to people and nations.[16] In the twenty-first century, however, critical infrastructure has become increasingly connected to the Internet. This increased connectivity is cause for global concern because much of the critical infrastructure that nations rely on for health, power, and security are susceptible to cyber threats. Attackers can now use virtual control systems to deliver physical threats or make virtual threats to physical infrastructure. This combination of virtual and physical threats is growing exponentially, especially as virtual connections to physical infrastructure via the Internet of Things become increasingly mainstream.[17] Such attacks will become more common and are likely to be more destructive. Beyond continuing to invest in the latest technology to help fend off these threats, having a strong CISR posture will minimize the impact of these attacks and make targets inherently less valuable from the adversary's perspective.

15. David Deming, "The Aqueducts and Water Supply of Ancient Rome," *Groundwate*r 58, no. 1 (January–February 2020): 153.

16. Chris Jensen, "What Is Critical Infrastructure and How Should We Protect It?," *Tenable* (blog), June 26, 2019, https://www.tenable.com/blog/what-is-critical-infrastructure-and-how-should-we-protect-it.

17. Jensen, "Critical Infrastructure?"

In 2019, security industry expert Pierre Bourgeix sounded the alarm regarding the increasing threats to critical infrastructure and the need to respond with enhanced security and resilience:

> We are on the precipice of a world that is completely connected. From smart cities to smart buildings, with the use of machine learning and deep learning, we are closer than ever to being completely converged. This connection is also our Achilles' heel that may lead to a disaster that we cannot possibly imagine. The need for our understanding of these threats from the lowest to the highest parts of our organizations is crucial. However, the time for education is growing short—either we get it or it is game over.[18]

From climate change to cyberattacks, infrastructure systems must operate in an increasingly uncertain future in which it is impossible to avoid or even predict all shocks and stresses. Therefore, it is essential for critical infrastructure stakeholders to prepare for the threats they can anticipate, and to be able to respond to the unexpected so that they can provide the essential services upon which society depends.[19] As Donald Rumsfeld, the then US defense secretary, famously said:

> There are known knowns; there are things we know we know. We also know there are known unknowns; we know there are some things we do not know. But there are also unknown unknowns—the ones we do not know we do not know. And if one looks throughout the history of free countries, it is the latter category that tends to be the difficult ones.[20]

Indeed, the known unknowns refer to the threats and risks of which critical infrastructure stakeholders are aware. Unknown unknowns are the risks that come from situations that are so unexpected that they would not be considered. To identify the future threats to critical infrastructure, stakeholders must try to envision them. It is certain that the years ahead will be marked by turbulence, fueled by unconventional conflict, fraught

18. Pierre Bourgeix, "Critical Infrastructure Security in a Converged and Interconnected World," Security InfoWatch (website), February 8, 2019, https://www.securityinfowatch.com/critical-infrastructure/article/21067817/critical-infrastructure-security-in-a-converged-and-interconnected-world.

19. Bourgeix, "Critical Infrastructure Security."

20. Donald H. Rumsfeld, "DoD News Briefing—Secretary Rumsfeld and Gen. Myers" (presentation, Pentagon, Washington, DC, February 12, 2002).

with natural and technological disasters, and complicated by the issues resulting from the increasing complexity of and societal dependence on critical infrastructure.

As this handbook stresses throughout all its chapters, enhanced CISR is essential to NATO's ability to function in crisis management, collective defense, and/or external operations. Securing a nation's critical infrastructure systems and making them more resilient against a wide range of current, emerging, and future risks is a complex and continuing challenge. In fact, it is one of the most difficult things a nation can do. Failing to achieve CISR goals and objectives, however, will reduce NATO's mission capability, and also adversely impact member states' collective societies because critical infrastructure is the foundation on which vital societal and economic functions depend. Therefore, one of the quintessential societal tasks necessary for maintaining national security, economic vitality, and public health and safety in a world filled with risk will be the continuing development, establishment, and maintenance of a demonstrably effective national CISR posture. Working together, sharing lessons to be learned, good practices, case studies, methods, tools, approaches, and experiences, and discovering the unknown unknowns will promote and enhance local, regional, national, and global CISR today and in the future.

— 2 —

Physical Threats to Critical Infrastructure

Malcolm Baker, Ronald Bearse, and Ray Mey

The September 11 (9/11) attacks on the United States demonstrated the vulnerability to the physical and kinetic threats posed by terrorists. These attacks also validated how determined, patient, and sophisticated terrorists have become in both planning and executing their operations. More than two decades later, the consequences of not securing critical infrastructure remain severe.

The basic nature of free societies greatly enables operations and tactics employed by terrorists, competitive states, and other malicious actors while hindering these societies' ability to predict, deter, mitigate the effects of, respond to, and recover from malevolent attacks against them. Therefore, it is imperative for member states of the North Atlantic Treaty Organization (NATO) to develop and sustain demonstrably effective critical infrastructure security and resilience (CISR) policies, plans, and procedures to reduce the risks to critical infrastructure posed by an ever-increasing list of credible threats, whether they are physical, cyber, or hybrid in nature. Nearly every day, threat actors commit cyberattacks against elements of critical infrastructure. Many of these cyberattacks receive a great deal of public and political attention because they harm NATO member states and partner nations, but physical attacks against critical infrastructure also remain dangerous threats, and they drive CISR planning and efforts on both sides of the Atlantic Ocean.

Most people have some level of personal experience with the COVID-19 virus and have witnessed how the pandemic, a biological event, has wreaked

havoc and caused damage across multiple critical infrastructure sectors in the United States, across NATO member states, and in the majority of the world's countries. Unfortunately, many people have also experienced firsthand some form of natural disaster, which ultimately causes billions of dollars in damage, adversely affects public health and safety, and perhaps even cripples or destroys some type of critical infrastructure. Many readers of this book also have an appreciation for the damage and destruction to critical infrastructure that occurred because of the 9/11 attacks. A less known but still disruptive event took place in 2016, when a monkey knocked out Kenya's entire power grid after falling onto a transformer at the Gitau Hydroelectric Power Station.[1] Such an unbelievable scenario—along with the examples of a terrorist attack, natural disaster, and biological event—illustrates that the current risk environment is very complex and uncertain as threats, vulnerabilities, and consequences have all evolved since the start of the twenty-first century.

This chapter provides a general overview of the current and emerging physical threats to critical infrastructure in a straightforward and thought-provoking manner. The chapter will proceed in three major sections. Each section represents one of the three basic categories of physical threats to critical infrastructure: (1) natural threats, (2) man-made threats, and (3) accidental or technical threats. Each of these sections will examine the nature of the threat to critical infrastructure, provide a few examples for each type of threat, and then examine relevant case studies to identify key insights and learning points to consider. The chapter will prioritize man-made threats, which is the longest section and contains multiple examples and case studies of the different types of threat in this category. The chapter will conclude with an overview of considerations for Allies and partners to enhance their CISR policies and practices to contend with the challenges emanating from physical threats to critical infrastructure.

Natural Threats

Natural events and disasters occur on a daily basis around the globe. This category includes such diverse natural threats as the effects of climate change, earthquakes, tsunamis, land shifting, volcanic eruptions, forest fires, hurricanes, floods, drought, and, in some circumstances, even time and animals. Natural threats pose a significant risk to critical infrastructure,

1. Jeff Postelwait, "Strangest Animal-Caused Power Outages," T&D World (website), May 27, 2021, https://www.tdworld.com/electric-utility-operations/media-gallery/21165503/strangest-animalcaused-power-outages?slide=10&id=21165503.

as the following examples clearly demonstrate. In 1993, a severe flood of the Missouri River threatened the safety of the Cooper nuclear power station in Nebraska. The Kobe earthquake in Japan in 1995 destroyed critical transportation, maritime, and chemical infrastructures. In 1998, tornadoes with wind speeds between 113 to 156 miles per hour hit the Davis-Besse nuclear power station in Ohio, and while this particular incident did not produce any long-term effects, it knocked out several critical systems, making the station more vulnerable to disaster. The widespread damage to energy and chemical infrastructures caused by Hurricanes Katrina and Rita in the United States in 2005 illustrates the threat posed by extreme weather. Events related to extreme weather, particularly lightning and storms, have historically posed the biggest threat to critical energy infrastructure.

Climate hazards like extreme weather events, higher temperatures, droughts, floods, wildfires, storms, rising sea levels, soil degradation, and acidifying oceans are intensifying and threatening infrastructure, health, and water and food security. Irreversible damage to ecosystems and habitats—degraded by air, soil, water, and marine pollution—will undermine the economic benefits they provide. Extreme weather events, many worsened by the accelerating sea level rise, will particularly affect urban coastal areas in South Asia, Southeast Asia, and the Western Hemisphere. Damage to communications, energy, and transportation infrastructure could affect military bases located in low-lying areas, inflict economic costs, and cause human displacement and loss of life.[2] Even what seem like unlikely events (such as a meteor hitting a major city and causing damage to multiple critical infrastructure sectors) are in the realm of the possible.

The frequency and costs of natural disasters are increasing, and they often pose a credible threat to specific elements of critical infrastructure located in the impact zone of the disaster. Due to the interconnectedness of critical infrastructure, sometimes a natural disaster in one area of a country will cause cascading effects across multiple critical infrastructure sectors, some as far away as the other side of the country or even in a neighboring country. The bottom line is that any threat assessment methodology must take into account the likelihood of damage or destruction to critical infrastructure resulting from natural hazards. Doing so enables both government and private owners and operators of critical infrastructure to better understand and manage the physical risks to critical infrastructure.

2. Daniel R. Coats, *Worldwide Threat Assessment of the US Intelligence Community: Statement for the Record, Senate Select Committee on Intelligence* (Washington, DC: Office of the Director of National Intelligence, January 29, 2019), 23, https://www.dni.gov/files/ODNI/documents/2019-ATA-SFR---SSCI.pdf.

Man-made Threats

Malevolent forces is a term used to describe the human actors who threaten to physically attack, disrupt, or destroy the normal operations of critical infrastructure. These forces include foreign and domestic terrorist organizations, elements working on behalf of a competitive foreign power, and individuals with malicious intent toward the government or a particular organization who are "super-empowered" by information, communications, and other technological advances.[3] This threat category includes not only war, terrorism, and hybrid warfare, but also rioting, financial crimes, economic espionage, and the possible use of chemical, biological, radiological, nuclear, and explosive (CBRNE) materials to inflict great harm. Traditional acts of sabotage and the intentional tampering, modification, or manipulation of physical systems and processes—such as physically moving a railway switching apparatus, opening and closing energy piping, and suppressing signals or alarms on critical nodes of energy and transportation infrastructure—remain a threat to critical infrastructure as well. Given this wide variety of threat actors and methods of attack, it is important to balance the likelihood and consequences specific threat vectors pose. See also the importance of plausibility outlined in chapter 7. This section will examine several types of man-made threats that are most grave and of greatest concern to NATO, and illustrate key concepts using case studies, beginning with the attack against the oil and gas facility at In Amenas, Algeria.

Case Study: In Amenas, Algeria

On January 16, 2013, approximately 32 heavily armed terrorists from an al-Qaeda affiliate attacked the In Amenas oil and gas facility in Algeria.[4] In an unprecedented, planned, and coordinated attack that spanned four days, it was one of the largest terrorist attacks ever conducted on an oil and gas facility. One of the largest gas developments in Algeria, the facility at In Amenas is operated as a joint venture between Statoil, British Petroleum, and Sonatrach, Algeria's national oil company. The vast facility covers over 2,700 square kilometers—an area almost equivalent to the size of Luxembourg—and is situated about 1,300 kilometers from the Algerian capital of Algiers and roughly 50 kilometers west of the

3. Toffler Associates, *Five Critical Threats to the Infrastructure of the Future: Leading Infrastructure Protection Experts Discuss Strategies for Protecting Your Enterprise* (Manchester, MA: Toffler Associates, 2008), 2, https://www.somanco.com/documents/Five%20Critical%20Infrastructure%20Threats.pdf.

4. Thomas Joscelyn and Bill Roggio, "Al Qaeda-linked Group Claims Credit for Kidnappings in Algeria," *Long War Journal* (website), January 16, 2013, https://www.longwarjournal.org/archives/2013/01/al_qaeda_commander_c.php.

Libyan border. Given its strategic importance, there was a multilayered security arrangement in place to protect the oil and gas facility.

The People's National Army and the Gendarmerie from the Algerian armed forces provided the outer layer of security for the facility, working separately and in support of one another. As in most other countries, the government was responsible for protecting the facility against terrorist attacks, including prevention, intelligence gathering, and border controls, including along the Algerian-Libyan border and in the desert zone surrounding the facility. In fact, the Algerian government had established a militarized zone around In Amenas with the army responsible for securing the wider desert area and the gendarmes providing security in the immediate vicinity of the facility. Regarding the inside the perimeter, Statoil, British Petroleum, and Sonatrach were responsible for inner security at In Amenas, and they provided an unarmed civilian guard force for this purpose. The internal guard force provided access control and other protective security measures as well as training and contingency planning to protect the facility's people and assets.

To initiate, the terrorists first attacked a bus carrying workers to the facility and then launched near-simultaneous assaults on the workers' living compound and oil and gas production area. Terrorists took control of the facility in just over 15 minutes. The situation entered a siege phase as terrorists searched for expatriate workers in the living and production areas. Although the attackers seized expatriate workers and corralled them together as hostages, they did not take any Algerian workers captive, or they allowed them to escape. The terrorists demanded the withdrawal of Algerian military forces from the vicinity and requested free and safe passage to Mali. The terrorists also demanded safe passage from the living compound to the production area buildings so that they could regroup together in one stronghold.

The Algerian military launched a helicopter strike against the attackers in the living compound, which prompted the terrorists to use the hostages as human shields. Subsequently, the terrorists tried to move the hostages in vehicles from the living compound to the production area. When the Algerian military interdicted the attackers' convoy, the terrorists returned fire and detonated improvised explosive devices (IED). Ultimately, the military succeeded in securing the living accommodations and resolving the situation, but at considerable cost in terms of human life. During the terrorists' assault and the military response, 40 people from 10 countries were killed along with 29 of the 32 terrorists.

The In Amenas attack captured global media attention given the nature of the attack and the type of critical infrastructure facility that the terrorists targeted. Among the many findings uncovered during the post-incident investigation, Statoil's final report concluded that the multilayered security arrangements failed to protect the living area and production areas from the terrorist attack because of an overreliance on the Algerian military to protect the facility. Furthermore, there was a lack of a strong security culture among the facility's senior executives, who did not believe that a credible threat existed and thus did not invest in security preparedness and resources to protect the facility. Following the release of the Statoil report, the In Amenas facility created a new security organization, implemented a security improvement program, introduced internal security governance structures, and developed and implemented a new security risk and threat assessment process. Underpinning all of these measures was a commitment to improve training and multiagency cooperation.

As the In Amenas attack illustrates, critical infrastructure stakeholders must deal with the problem of dynamic and evolving adversaries. Since these threat groups typically have rapid planning cycles, are constantly adapting their strategies, and can adopt new tactics to adjust to updated security measures, NATO member states and partner nations must consider how to stay ahead of such adversaries.[5] One answer to this vital question is that Allies and partners must keenly focus on ensuring their critical infrastructure systems are made more secure and resilient against the current and emerging threats and hazards they are most likely to encounter. While the list of threats posed by malevolent actors is quite extensive, there are several types of these man-made threats that are of greatest concern, which this section will now explore in detail.

Insider Threat

The insider threat is a major man-made physical threat to critical infrastructure that is often overlooked and underestimated despite the growing number of attacks perpetrated in this manner. In fact, insiders pose the greatest threat—especially if they are working with a foreign state or other high-level threat groups—because of their detailed knowledge of system operations and security practices. Insiders are often company employees or suppliers, and they can act either as the main conspirators or as accomplices and informants

5. Toffler Associates, *Five Critical Threats*, 2–3.

to outside actors. Unlike external actors, who can only gain access to elements of critical infrastructure by means of violent acts or subterfuge, insiders have undisputed advantages. Given their access and placement to observe a facility's processes undisturbed over time, the knowledge insiders possess or the ease with which they can acquire it is of high value to a wide range of potential malevolent actors.

For the foreseeable future, the list of possible insider threats to critical infrastructure includes disgruntled employees seeking revenge, hackers testing their skills, criminals seeking financial gain, foreign intelligence operatives seeking sensitive government or industrial information, and terrorist groups or hostile nations conducting attacks on vital services (such as electric power, transportation, energy, or telecommunications systems). Along with the physical threat insiders pose to critical infrastructure, the insider threat also affects the cyber domain where the anonymity of cyber threat groups makes it difficult to identify those responsible for an intrusion or to ascertain their intentions. With these factors in mind, methodologies for conducting risk assessments of specific systems and vulnerable locations within critical infrastructure should include each employee's role within the system as well as all visitors to the site or facility. For instance, when analyzing the threat of a person-borne IED (PBIED) used to attack an aircraft, personnel conducting an assessment should consider, separately, both a PBIED smuggled onto the aircraft by a passenger and one brought on board by crewmembers and/or other airline employees. Implementation of effective personnel hiring and vetting procedures is a key preventive measure owners and operators of critical infrastructure can take to enhance organizational security.

The insider threat has existed since the earliest civilizations, with stories of insider threats present in nearly all cultures. Examples from American history include everything from Benedict Arnold's betrayal to recent unauthorized and devastating disclosures of classified information. Along this spectrum, there is a common narrative: trusted and capable people, often facing enormous life challenges, exploit their access and trusted status to betray their organization and ultimately harm the nation. Given the resources that foreign adversaries are dedicating to exploit or coopt insiders within organizations they seek to penetrate, insider

threats will be an enduring part of the threat and risk landscape for most critical infrastructure entities for years to come.[6]

Since insider threats are inherently a human problem, critical infrastructure stakeholders ultimately need to adopt human solutions to counter them. See also chapter 3 for its discussion of the importance of human capital in defending against cyber threats. While technology can help organizations understand their employees' activities in the cyber and virtual realms, the most powerful weapons critical infrastructure owners and operators have to counter insider threats are the personnel who comprise the organization's work force. To help mitigate these threats, organizations need to identify abnormal or suspicious activity among their employees and then respond accordingly in ways that foster trust and leverage the workforce as a partner.[7] Insider threats are an increasingly important threat vector to critical infrastructure in the context of broader security risks as well as those related to supply chains and cybersecurity. While insider threats can cause damage through economic espionage, sabotage, workplace violence, fraud, and other misuse of corporate resources, the various threat activities can include deliberate actions by insiders working with foreign intelligence services or other actions by insiders with malicious or criminal motives. Finally, insiders can unknowingly commit mild to severe security breaches due to simple carelessness, utter negligence, or complete disregard for simple security rules and procedures.

In summary, the insider threat is very real. It represents a significant risk to critical infrastructure assets and is a hazard that critical infrastructure stakeholders must consider in the process of assessing and managing overall risk.

CBRNE Threat

The malevolent use of CBRNE materials is another type of man-made physical threat to destroy or damage critical infrastructure. In the hands of malicious actors, these materials pose a significant threat to the populations, infrastructure, economies, and security of NATO member states. Therefore, securing the public, emergency responders, agricultural

6. National Counterintelligence and Security Center (NCSC), *Insider Threat Mitigation for U.S. Critical Infrastructure Entities: Guidelines from an Intelligence Perspective* (Washington, DC: Office of the Director of National Intelligence, 2021), 2–4, https://www.dni.gov/files/NCSC/documents/news/20210319-Insider-Threat-Mitigation-for-US-Critical-Infrastru-March-2021.pdf.

7. NCSC, *Insider Threat Mitigation*, 4.

resources, and critical infrastructure against these threats must always be a priority. This section examines each element of the CBRNE acronym—chemical, biological, radiological, nuclear, and high-yield explosives—to understand the nature of the threat and outline potential measures to mitigate these threats and enhance CISR policy and practices.

The first category in CBRNE, chemical materials, poses a daunting challenge because threat actors can access these materials with relative ease. NATO member states and partner nations manufacture and store chemicals in abundance due to their role as an integral part of modern life. Malevolent actors such as terrorists, extremists, and saboteurs can use chemicals common in industrialized nations to create improvised explosives, incendiaries, chemical agents, or even an improvised chemical weapon. Some of the more common types of chemicals that threat actors could use in improvised chemical weapons include acids, ammonia, benzene, chlorine, and propane. These chemicals are manufactured or stored in various locations, such as manufacturing plants, industrial facilities, gas stations, and research laboratories. Since chemicals are routinely transported using other critical infrastructure sectors—such as rail networks, waterways, roads, and aircraft—they are easy targets for sabotage and even more difficult to protect.

The use of man-made sarin gas to attack the Tokyo subway system in 1995 is a sobering reminder of the vulnerabilities and impact associated with the weaponization of dangerous chemicals. See chapter 7 for more detail on this attack. It is also a vivid reminder of the potential dangers associated with the motivations of malevolent groups, whether foreign or domestic terrorists, proxy forces, elements working on behalf of an adversarial foreign power, or the lone wolf attacker with malicious intent toward a government, organization, or company.

The next category among CBRNE materials, biological threats, involves the intentional and accidental release or dissemination of biological agents like bacteria, viruses, fungi, or toxins—spread through the air, water, or in food—to cause illness, death, and panic. Although the possibility of attack is relatively low because of the challenges to cultivate, weaponize, and deploy biological agents, acts of bioterrorism can still harm critical infrastructure. Recent outbreaks of Escherichia coli (E. coli) highlight how easily a biological attack on critical agricultural infrastructure could undermine consumer confidence and cause a host of public health problems.

In a recent example, a meat distributing company in the United States had to recall more than 14 tons of beef due to a potential contamination with E. coli.[8] This example demonstrates that agriculture and food systems are vulnerable to diseases, pests, or poisonous agents that occur naturally, are unintentionally introduced, or intentionally delivered by threat actors. Certain insects, for instance, can destroy forests or agricultural zones, causing billions of dollars in economic damage. Consequently, research facilities studying invasive species must manage the risk of their uncontrolled or unauthorized release.

The current global pandemic attributed to the emergence and rapid spread of the COVID-19 virus in Wuhan, China, represents a physical biological threat vector. Biological agents have the potential to kill millions of people across the world with deleterious impacts on critical infrastructure, not to mention the trillions of dollars required to respond to and recover from the socioeconomic damage inflicted.[9] The impact of COVID-19, as well as the possible weaponization and release of new viruses or other biological agents, demands that nations strengthen their ability to detect, deter, destroy, or prepare for, mitigate the effects of, respond to and recover eventually from the release of such toxins.

Many "lessons-to-be-learned and applied" will be discussed, defined, and shared across the globe, but the unfortunate fact remains that lessons that should be learned often are not. Consequently, when key lessons are disregarded or not learned at an institutional, collective level, similar events that occur in the future are susceptible to unfolding in much the same way with the same results. Together, this wide spectrum of possible biological threats cannot simply be disregarded, even if some are more likely to occur than others. As chapter 13 highlights in its robust examination of assessing and managing risk, once governments identify which infrastructures are truly critical, all hazards that pose a significant risk to them should be clearly defined, assessed, and considered in defining and managing risk.

The use of radiological and nuclear weapons, the third and fourth categories of threat within the CBRNE framework, represents another dangerous physical threat against critical infrastructure. The bad news on this front is that increasingly well-organized and well-funded terrorist organizations,

8. David K. Li, "Over 28,000 Pounds of Ground Beef from Oregon Recalled over Possible E. Coli Contamination," *NBC News* (website), January 7, 2022, https://www.nbcnews.com/news/us-news/28000-pounds-ground-beef-oregon-recalled-possible-e-coli-contamination-rcna11414.

9. "WHO Coronavirus (COVID-19) Dashboard," World Health Organization (website), accessed January 24, 2022, https://covid19.who.int/.

which now have easy access to the expertise needed to build a bomb, have declared their intent to seek the materials necessary for building and using weapons of mass destruction. Although the probability of a terrorist group building even a crude clear device or sabotaging a nuclear power plant is low, the international community cannot afford to be complacent. The good news is that international political leaders largely share the same threat perception of nuclear terrorism and increasingly are cooperating on this key policy issue to mitigate the risks emanating from it. According to a report by the Arms Control Association and the Fissile Materials Working Group that compiles data from 2010–16, countries made sizable progress to curb nuclear terrorism and strengthen and improve nuclear security.[10] These voluntary national commitments represent some of the most tangible and innovative nuclear security protocols in existence, and countries are actively practicing and implementing them across the globe. There is no credible evidence that any terrorist group has succeeded in obtaining the necessary multi-kilogram critical mass amounts of weapons-grade plutonium required to make a nuclear weapon. Even if a group did possess a few kilograms of fissile material, a crude terrorist-built design would require far more material and pose a huge technical challenge. At the same time, security practitioners cannot simply dismiss such a possibility.

The prospect of terrorists or other malevolent actors building and detonating a so-called dirty bomb—otherwise known as a radiological dispersal device (RDD)—is considerably higher because constructing an RDD is no more complicated than building an IED. An RDD combines conventional explosives with materials containing radioactive isotopes, some of which can be found in products such as equipment to treat cancer, photocopiers, watches, and even household smoke detectors containing the radioactive isotope Americium 241. The detonation of an RDD would likely have a greater psychological impact on the affected population compared to the physical damage it would cause. Given the feasibility of an RDD attack, critical infrastructure stakeholders must consider to what extent response forces are prepared to mitigate the physical and psychological effects of the explosion, and what impacts such an event would have on facilities and systems.

Beyond the damage to property and critical infrastructure, the malevolent use of CBRNE materials has the potential for long-lasting effects on the mental health and psyche of the citizenry and first responders involved

10. Sara Z. Kutchesfahani and Kelsey Davenport, "Why Countries Still Must Prioritize Action to Curb Nuclear Terrorism," Bulletin of Atomic Scientists (website), August 3, 2018, https://thebulletin.org/2018/08/why-countries-still-must-prioritize-action-to-curb-nuclear-terrorism/.

in such an incident. For instance, the physical and psychological trauma of the bombing of the Afred P. Murrah Federal Building in Oklahoma City and the 9/11 attacks in the United States led to post-traumatic stress disorder (PTSD) and increased the use of and addiction to anxiety drugs. Among the nearly 37,000 members of the World Trade Center Health Registry, roughly 14 percent suffered from PTSD and 15 percent from depression more than a decade after the attacks, and studies indicate approximately 10 percent of first responders were still dealing with PTSD more than a decade later.[11] Similarly, the adverse mental and psychological effects of the COVID-19 pandemic show the heavy toll on societies that could result from physical attacks impacting critical infrastructure. Adversaries have learned from these and other catastrophic events, and are certainly factoring in these types of effects into their prosecution of physical attacks and use of hybrid threat vectors to divide Western societies.

The final threat category in the CBRNE family, explosives, is the most common type of man-made physical threat to critical infrastructure. This category involves the use of explosives such as IEDs, rockets, and grenades to cause damage, disruption, or destruction of critical infrastructure. Beyond conventional explosives, this category also includes tools as simple as lighters and matches, crude incendiary devices, vehicle-borne IEDs, and homemade explosive vests. In this sense, something as simple as a petrol bomb or Molotov cocktail, coupled with a threat actor's malicious intent, can create disastrous consequences, damaging or destroying a critical infrastructure's physical facilities and disrupting or halting the essential services they provide. Therefore, security measures for critical infrastructure need to consider more than just protection against conventional IEDs. Some potential protective measures include: timely, accurate information; advanced automated surveillance (such as CCTV), detection devices for indoor and outdoor use, prioritization of threats by installed systems (to avoid constant false alarms), communications interoperability, and, of growing importance, the resilience of the organization or facility following an attack.

Drone Threat

Another type of man-made physical threat to critical infrastructure is the malevolent use of drones. Drones are any vehicle that can travel autonomously; though commonly treated as synonyms, drones and unmanned aerial vehicles

11. Jonathan Strum, "The Effect 9/11 Had on Mental Health in America," Recovery Village (website), September 10, 2019, https://www.therecoveryvillage.com/mental-health/news/9-11-effect-mental-health/.

(UAV) are not entirely the same. An autonomous aerial drone does not require human intervention but can fly using its onboard autopilot, computer, and suite of sensors. Although a UAV does not require a human pilot or crewmembers to be on board the aircraft for it to fly, human operators typically pilot the UAV remotely from the ground.[12] The UAV is a component of an unmanned aircraft system, which includes not just the UAV but also everything required for it to function: the ground-based control module, its global positioning system module, its system of communications with the UAV, and the person on the ground controlling the UAV.[13] Despite these nuanced differences, this section will use the generic term *drone* to refer to the category of physical threat posed by the entire spectrum of fully and semi-autonomous aerial systems.

When operated by tech-savvy lone-wolf actors, terrorist groups, domestic and transnational criminal organizations, or subservient proxies operating in combat locations, disputed territories, or even terrain overlooking vulnerable critical infrastructure nodes, drones pose a significant threat to critical infrastructure. Therefore, critical infrastructure security risk assessment and management processes need to incorporate this physical threat vector into overall CISR posture. Depending on its power and size, a drone can transport contraband, chemicals, or other explosives or weaponized payloads. These systems are also capable of silently monitoring a large area from the sky for nefarious or strategic intelligence purposes, and can even be used to perform cyberattacks involving the theft of trade secrets, technologies, or sensitive information affecting critical infrastructure. There are numerous examples of malicious activities using drones to target critical infrastructure and other sensitive sites or personnel. In France, seven nuclear power plants confirmed unauthorized drone flights above their facilities in 2014, while Greenpeace activists intentionally crashed a drone into a French nuclear power plant near Lyon in 2018.[14] These examples show that drone attacks are an emerging physical threat vector to carry out attacks against critical infrastructure—a trend that is likely to increase in the years ahead.

Available for as little as $1,200, top-selling quadcopters can carry items weighing up to one kilogram. Threat actors could exploit the accessibility and

12. "Unmanned Aerial Vehicles (UAV), Unmanned Aerial Systems (UAS), and Autonomous Drones: What's the Difference?," MissionGo (website), accessed on January 24, 2022, https://www.missiongo.io/unmanned-aerial-vehicles-uav-unmanned-aerial-systems-uas-and-autonomous-drones-whats-the-difference/.

13. "What's the Difference between Drones, UAV, and UAS? Definitions and Terms," Pilot Institute (website), March 22, 2020, https://pilotinstitute.com/drones-vs-uav-vs-uas/.

14. Clay Taylor, "No Trivial Offense: Drones Disrupt Nuclear Power Plant Security," *Dedrone* (blog), August 19, 2020, https://blog.dedrone.com/en/no-trivial-offense-drones-disrupt-the-security-of-nuclear-power-plants.

mobility of quadcopters to transport a camera to film bridges, government buildings, stadiums, or motorcades, looking for security flaws or even conducting attacks using lightweight explosive devices. Due to their physical and operational characteristics, many types of drones can evade detection and create challenges for the owners and operators of critical infrastructure. Given the continued accessibility and versatility drones provide, malevolent actors will likely use drones in any of the following capacities. First, they can use drones for reconnaissance of critical facilities to gather intelligence on the site layout, guard movements, or other information that could help in carrying out a physical attack. Second, threat groups can use drones to drop explosives intended to damage critical or sensitive infrastructure in transportation hubs and other areas of public gathering. Third, drones can deliver weapons or other materials for use in an attack. Finally, drones can provide air support and overwatch to support a ground attack. To demonstrate the lethality of drone attacks against critical infrastructure, it is appropriate now to examine the case study of the Saudi Aramco attack.

On September 14, 2019, Houthi rebels based in Yemen claimed responsibility for using a combination of drones and cruise missiles to attack two strategic oil facilities in Saudi Arabia operated by the state-owned company, Saudi Aramco. Reports indicated 19 points of impact at the oil processing plant at Abqaiq—the largest in the world—and the Khurais oilfield, which started fires that caused considerable destruction and disruption to the facility operations.[15] Although Houthi rebels took credit for the attack, some Western intelligence agencies dispute this claim. The advanced drone technology used in these attacks represented a dramatic escalation in the conflict and in the rebels' capability to mount such attacks, suggesting that this attack was a state-sponsored or state-enabled act of terrorism in the region.

Despite the uncertainty regarding the specific threat actors behind the attack, the damage to the very heart of Saudi Arabia's oil infrastructure was indisputable. The attacks damaged more than half of the depressurization units and three of the stabilization columns at the Abqaiq oil facilities, decreasing Saudi Arabia's daily oil production by some five million barrels per day over 10 days.[16] This loss, which represents about five percent of global

15. "Saudi Oil Attacks: Drones and Missiles Launched from Iran -US," BBC (website), September 17, 2019, https://www.bbc.com/news/world-middle-east-49733558.

16. Katie McQueen and James Leach, "A Year after Abqaiq Attacks Saudi Aramco Still Seen Vulnerable," S&P Global Commodity Insights (website), October 13, 2020, https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/oil/101320-feature-a-year-after-abqaiq-attacks-saudi-aramco-still-seen-vulnerable.

oil production, caused global oil prices to jump to a 20-year high as a result.[17] The security at the Abqaiq plant was deemed extremely high, so the destruction caused by the attacks increased concerns of a broader regional conflict and heightened global tensions that similar or further attacks against critical infrastructure could intensify.[18] Perhaps most alarming is that the attacks against the Abqaiq facility, which some experts refer to as the crown jewel of the Saudi oil infrastructure and the beating heart of the global oil system, brought into question Saudi Arabia's valued reputation as a dependable, stable oil supplier to the global market.[19]

The considerable post-incident learning included the balance of threat, risk, and harm when considering how to make critical infrastructure more secure. The nature of oil, gas, and energy production illustrates the extensive interdependencies with other infrastructure sectors, economic futures, and global markets. Operators of critical infrastructure provide services that are difficult to protect from the air without support by the military or security forces. Mitigating threats to oil and gas facilities is quite different from countering threats to other forms of critical infrastructure, including modes of transportation such as aviation and railways. See chapters 6 and 7 for further insight into securing against threats to civil aviation and rail. These case studies effectively demonstrate that security practitioners cannot take security arrangements and efforts to mitigate threats in a specific sector and then apply them uniformly across all critical infrastructure sectors. It is necessary to have a balance based on threat, plausibility, credibility, and proportionality. See chapters 7 and 13 for a discussion of these concepts as part of a broader strategic risk assessment framework.

This brief survey of how malevolent forces have used drones in the past begs the question of how long it will be before threat actors escalate their tactics and use drones more extensively for targeting critical infrastructure. One such danger is drone swarming, which is a tactic that employs several drones simultaneously to coordinate operations and accomplish tasks that a single drone cannot. In this configuration, each drone may perform a similar function or have unique, specialized roles, such as gathering information,

17. Laila Kearney, "Oil Jumps Nearly 15% in Record Trading after Attack on Saudi Facilities," Reuters (website), September 15, 2019, https://www.reuters.com/article/us-global-oil/oil-prices-surge-15-after-attack-on-saudi-facilities-hits-global-supply-idUSKBN1W00UG.

18. Michael Safi and Graeme Wearden, "Everything You Need to Know about the Saudi Arabia Oil Attacks, *Guardian* (website), September 16, 2019, https://www.theguardian.com/world/2019/sep/16/saudi-arabia-oil-attacks-everything-you-need-to-know.

19. McQueen and Leach, "A Year after Abqaiq Attacks."

carrying weapons or explosives, or relaying communications. Drone swarming requires advanced capabilities, including the ability for individual drones to maintain spacing, avoid in-flight collision, and predict where other drones in the swarm will be at a given time.[20] To reach this level of sophisticated capability, drone swarms typically rely on real-time sensing, artificial intelligence, computer vision algorithms, and communications based on radio frequency, cellular, or satellite communications. By using a single ground control station to control the drone swarm, threat actors could simplify deployment and equipment requirements, and effectively operate the drones autonomously so that an operator does not have to control multiple drones in real time.[21] Without being too far-fetched, it is possible to imagine drone swarms capable of maneuvering against active defense measures, identifying targets, and delivering payloads as diverse as chemical agents over population centers or explosives against targets like oil fields or water treatment facilities.[22]

Examining the threat drones and drone swarms pose to critical infrastructure highlights the need for NATO member states and partner nations to take vital steps to enhance their CISR posture. In one such example, NATO's Communications and Information Agency has developed a low-cost prototype solution for rapidly detecting, identifying, and localizing small drones that may pose a threat. This prototype, known as the ARTEMIS system, employs electromagnetic waves to identify drones and advanced techniques to detect and classify radio frequency signals that the drones are using. The equipment has been successful in open field testing, with very promising results for use in mitigating the threat posed by commercial drones.

Threat of Precision Strike Weapons

Although rarely considered in homeland security assessments, precision attacks using missiles—traditionally thought of as conventional weapons—pose a threat to principal critical infrastructure systems across NATO. Precision missiles can engage targets at extended ranges, from 100 yards to thousands of miles. While the military already employs certain measures

20. "Drone Swarm Technology," Unmanned Systems Technology (website), accessed on February 11, 2022, https://www.unmannedsystemstechnology.com/expo/drone-swarm-technology/.

21. "Drone Swarm Technology."

22. Zachary Kallenborn and Philipp C. Bleek, "Drones of Mass Destruction: Drone Swarms and the Future of Nuclear, Chemical, and Biological Weapons," War on the Rocks (website), February 14, 2019, https://warontherocks.com/2019/02/drones-of-mass-destruction-drone-swarms-and-the-future-of-nuclear-chemical-and-biological-weapons/.

to thwart such stealthy attacks abroad and defend key military installations, other nodes of critical infrastructure are poorly defended or not defended at all, making them relatively soft targets. With an ever-increasing potential for terrorists to procure missile technologies and weapons, precision missile strikes could represent an enduring threat from both terrorists and rogue states. To understand the numerous precision weapons that could threaten critical infrastructure across NATO member states, this section will examine the threats posed by short-range and long-range systems.

In the category of short-range systems, the most common weapon is the man-portable air defense system (MANPADS), which was originally developed to defend against military aircraft. Since MANPADS have precision strike capabilities, are globally available, and come in a variety of configurations, terrorists have traditionally used them to target passenger aircraft. See chapter 6 for examples of terrorist attacks against civil aviation using MANPADS. Given their relatively small size—some systems weigh only 35 pounds and measure six feet long—MANPADS are easy to conceal and transport, making them ideal weapons of choice for threat actors looking to target vulnerable points at ground facilities, such as power plants or oil and gas facilities. Similarly, anti-tank guided missiles (ATGMs) are readily accessible on the weapons black market and equally easy to hide and transport. Malevolent actors could quite easily store ATGMs, along with their guidance system and ammunition, in a car trunk and use them to target any number of critical infrastructure nodes, such as major financial facilities, water treatment plants, and even primary government buildings.

Among long-range precision weapons systems, the threat cruise missiles pose to critical infrastructure is also growing. While relatively few nations have land-attack cruise missiles, many have anti-ship cruise missiles. Although the primary function of these systems is to target ships at sea, malevolent forces could also modify them to target critical infrastructure or simply use them as weapons of terror by launching them indiscriminately at populated areas. In the same way, ballistic missiles are becoming a more prevalent threat. Although few nations possess intercontinental ballistic missiles, many do have ballistic missiles capable of operating at shorter ranges. Malicious actors could smuggle these missiles on cargo ships and transport them globally for use against critical infrastructure targets or population centers in Alliance member states and partner nations.

Electromagnetic Pulse Threat

Another physical threat vector is an electromagnetic pulse (EMP) event and the catastrophic effects such an event would produce. An EMP event is a short burst of high-energy electromagnetism, which under certain circumstances can disrupt or destroy electrical and electronic capability. Since much of modern life depends on electricity and electronics, a widespread EMP attack could cause a major catastrophe. When interacting with the earth's magnetic field, these powerful pulses have the ability to damage electronic and electrical equipment like computers, cell phones, transformers, transmission lines, and the broader critical communications infrastructure. See chapter 9 for an overview of the communications network as a lifeline sector. Even worse, the design of many electrical grids means that damage to certain electrical systems, the lifeblood of any modern society, could cause failures throughout a number of systems—including banking, energy, transportation, food production and delivery, water, emergency services, and cyberspace—across an entire country or region.

Extreme electromagnetic incidents caused by an intentional, man-made EMP attack or a naturally occurring geomagnetic disturbance (also called space weather) could cause significant damage to critical infrastructure sectors, such as the electrical grid, communications and transportation, networks, and water and wastewater systems.[23] Cascading effects are likely to follow, with the impacts of an attack initially compromising one or more elements of critical infrastructure, then spilling over into additional sectors, and expanding beyond the initial geographic regions. Intentional attacks using high-altitude nuclear detonations, specialized conventional munitions, or nonnuclear directed energy devices can lead to EMP events. Depending upon the specific characteristics of the weapon and the attack profile employed, the effects of an EMP event can vary in scale from local to regional to continental.[24]

Governments all over the world, and particularly their respective military forces, have been dealing with EMP threat assessment since the early years of the Cold War. This process is a tough assignment because the required technical information is subject to change—especially for a nonnuclear EMP event—and the motivational aspects stemming from political concerns and

23. National Coordinating Center for Communications, *Electromagnetic Pulse (EMP) Protection and Resilience Guidelines for Critical Infrastructure and Equipment* (Arlington, VA: National Cybersecurity and Communications Integration Center, 2019), 2, https://www.cisa.gov/sites/default/files/publications/19_0307_CISA_EMP-Protection-Resilience-Guidelines.pdf.

24. National Coordinating Center for Communications, *Electromagnetic Pulse (EMP)*, 2.

the economic situation are difficult to evaluate and predict. Defense analyst Peter Pry, who formerly chaired the US Congressional EMP Commission, points out that detonating a nuclear weapon roughly 200 miles above the United States could cause an EMP event covering almost the entire North American continent, and malevolent actors or groups could build and employ an EMP generator powerful enough to cripple a major metropolitan area.[25] Whether resulting from an attack or a natural event, the risks emanating from an EMP event are considerable and they merit continued attention and action by critical infrastructure stakeholders across NATO countries.

Accidents and Technical Threats

The final physical threat category this chapter examines, accidents and technical threats, includes a host of disasters due to acts of omission or commission, technical or infrastructural failures of various kinds, and situations or events that result directly from human error. Among these threats are technical failures to power grids and water treatment facilities, faulty safety systems, aging infrastructure that causes a water main to rupture or hazardous materials to release, and even monkeys (or other things) falling into places they should not. Together, the consequences of these threats range from mild to severe.

Examples of these types of incidents include radiation leaks at nuclear power plants, toxic gas emissions from chemical plants, lethal emissions from fertilizer production plants, and human death and environmental destruction caused by oil and gas platform catastrophes. The release of deadly radioactive materials following the 1986 Chernobyl and 2011 Fukushima nuclear plant accidents are two specific examples in recent memory. The catastrophic effects of these two incidents also highlight the devastating potential use of these materials in the hands of malevolent actors. With the growing use of and reliance on nuclear energy, theft of nuclear materials and terrorist motivations to attack and release radiological materials are realities that demand robust security, safety support, and regulatory oversight to counter.

The pace with which modern economies have become inextricably linked over the past two decades, especially through the great strides made in information and communication technology, has exposed the societies

25. Edd Gent, "US Air Force Is Guarding against Electromagnetic Pulse Attacks. Should We Worry?" LiveScience (website), March 11, 2021, https://www.livescience.com/air-force-emp-attacks-protection.html.

and critical infrastructure of NATO member states to a set of unprecedented threats and vulnerabilities. Even simple wear and tear poses a constant challenge to critical infrastructure. Although critical infrastructure does not fail because of advanced age or lagging maintenance alone, aging assets can degrade performance and lead to functional obsolescence, both of which increase the risk of failure. Incidentally, accidents attributed to the use of farm equipment, such as using a backhoe to dig in the ground, account for more disruption of communications systems per day than any other physical event. What is unfortunate about the threats from accidents and technical failures is that any one of them has the potential to produce unacceptable losses in terms of human casualties, property destruction, economic effects, or public morale and confidence. This stark reality is precisely why all three categories of physical threat—natural disasters, man-made events, and accidents and technical failures—must be taken into consideration when determining risk to specific elements of critical infrastructure in any sector.

More to Consider: Threats to Port Facilities

Maritime transport—comprised of dry bulk ships, container ships, and oil tankers—accounts for 80 percent of the world's trade volume and 70 percent of its trade value, making major sea ports the hubs of all global economic activity.[26] Many critical infrastructure facilities in NATO member states are located along coastlines, to include oil and gas terminals, desalination plants, and nuclear plants that require large amounts of water for the cooling process. These major ports present a target rich environment for malevolent actors; the right target, in the right location, at the right time using the right threat vector could cause enormous damage to port facilities and to the societies and economies they support. Security and resilience planners as well as owners and operators of port facilities have much to consider when thinking about the physical threats to major ports. To get a better sense of the challenges of protecting these vital facilities, stakeholders in sea ports must consider the port from a geospatial perspective, which requires an examination of the water around the port—both the surface level and under the water—as well as the port facilities on land and the airspace above them.

Regarding the physical threats emanating from the water surrounding port facilities, several terrorist organizations have demonstrated their intent to develop capabilities to conduct kinetic strikes from the sea, in particular

26. Myrto Kalouptsidi, "The Role of Shipping in World Trade," Econofact (website), June 9, 2021, https://econofact.org/the-role-of-shipping-in-world-trade.

those that combine the use of scuba divers and explosives.[27] There are two predominant methods divers employ to attack ports: they either attach some sort of a mine to a craft adjacent to the port or they emplace a mine on the port's channel bed. Divers may find it challenging to transport and emplace a large quantity of explosives that would be required to have a significant impact when attacking most large vessels. Nevertheless, a strategically placed charge could cause significant damage and difficulties for cruise ships, naval vessels, or a single-hulled tanker filled with components that support a wide variety of critical infrastructure. This idea is not far-fetched, as an example from December 2014 illustrates. Then, Egyptian security forces arrested members of a terrorist cell in Sinai for attempting to carry out an attack against Egyptian warships. During the arrest, security forces found scuba diving gear among the members of the terrorist cell, which was already complicit in previous attacks using small arms weapons and missiles to conduct ambushes and hijackings against Egyptian naval vessels.[28]

Similarly, waterside port security must account for attacks from a number of platforms malevolent actors could use, including fast boats, jet skis, shipping containers, and remotely piloted boats. When considering which types of attacks are plausible, there are two specific examples that are especially helpful. The first attack vector, pursued by terrorist groups like al-Qaeda, is to smuggle an IED into a ship-borne container for detonation against targets far from the original launch site as a form of power projection for terrorist groups.[29]

A second potential attack mode is to create a swarm that combines suicide boat drones and piloted suicide boats—effectively a maritime version of the aerial drones and drone swarms described in the previous section. This attack vector is similar to how the Japanese Imperial Navy and Army used *Shinyo* suicide boats, capable of carrying more than 500 pounds of explosives and traveling nearly 30 miles per hour, to cause significant damage to US ships during World War II.[30] In the modern context, Houthi rebels used a remote-controlled boat loaded with explosives—likely provided by Iran's Revolutionary Guard Corps—to attack the Royal

27. Meghan Curran, *Soft Targets & Black Market: Terrorist Activities in the Maritime Domain* (Broomfield, CO: One Earth Future, 2019), 10–12.

28. Norman Cigar, *The Jihadist Maritime Strategy: Waging a Guerrilla War at Sea* (Quantico, VA: Marine Corps University, 2017), 17–18, https://www.usmcu.edu/Portals/218/MES/Monographs/MESM_8_MAY_2017_lo.pdf?ver=2018-10-16-110147-393.

29. Cigar, *Jihadist Maritime Strategy*, 28.

30. Bob Hackett and Sander Kingsepp, "Shinyo! Battle Histories of Japan's Explosive Moorboats," Combined Fleet (website), November 26, 2011, http://www.combinedfleet.com/ShinyoEMB.htm.

Saudi Navy's frigate *Al Madinah* in January 2017.[31] While this attack vector is similar to the suicide boat attack against the USS *Cole* in 2000, the capability to conduct unmanned attacks represents a more dangerous threat because malevolent forces can be bolder and more lethal when their own lives are not directly at stake. With respect to physical threats to land-based port facilities, the vehicle-borne IED is, for good reason, still the weapon of choice for most terrorists. These mobile weapons can pack large amounts of explosives into a vehicle, and the bigger the vehicle, the bigger the explosion. Not only can threat actors drive a vehicle-borne IED directly to the target location, these attacks can also be difficult to detect and prevent.

As this chapter discussed earlier, malevolent forces can take advantage of the availability and mobility of drones to target land-based port facilities from above with the goal of crippling key machines or neutralizing the work force either temporarily or permanently. Such attacks can be devastating enough during routine port operations and daily business. They can be even more damaging when NATO member states are in the process of mobilizing military forces and equipment in support of collective defense, crisis management, or other external operations. See chapter 4 for a detailed discussion of the importance of maritime transportation to NATO operations and the challenge posed by hybrid threats. The difficulty of building and maintaining secure, resilient ports is certainly formidable. However, significant investments of thought and resources, the development of advanced situational awareness capabilities for a common operating picture, and the deployment of effective threat countermeasures and state-of-the-art protection systems are helping to secure major ports. This accomplishment represents some of the very best CISR efforts in a number of nations.

Increasing Sophistication and Outsourcing of Physical Threats

Since the 9/11 attacks, the ability to monitor, detect, and defend against a wide range of physical threats to critical infrastructure has increased considerably, as demonstrated by the vast number of video-surveillance cameras—and their attendant software and human controllers—surrounding almost every critical infrastructure site or facility in many Western nations.

31. Sam LaGrone, "Navy: Saudi Frigate Attacked by Unmanned Bomb Boat, Likely Iranian," *USNI News* (website), February 20, 2017, https://news.usni.org/2017/02/20/navy-saudi-frigate-attacked-unmanned-bomb-boat-likely-iranian.

Unfortunately, as the defenders have become more sophisticated, so too have the attackers and their capabilities.[32]

Since private-sector entities own and operate much of the critical infrastructure in the West, expensive security measures inevitably compete against an array of economic considerations, creating a reality in which security alone is never the deciding factor. This dynamic creates two unique vulnerabilities. The first, resource disparity, is based on the fact that securing critical infrastructure against physical threats can be an expensive venture that requires the allocation of significant resources but must contend with other competing priorities as well. Often there are not sufficient resources to address all potential security challenges, or there is not a strong enough rationale to justify spending money on security instead of improving business efficiency or enhancing operations. Large companies and organizations are more likely than small companies to be able to afford such expenditures. The second vulnerability, outsourcing complexity, highlights the tendency for companies and organizations to focus on core competencies, and outsource all else to outside providers, to include areas such as transportation, utilities, healthcare, and financial-service providers. Quite often, physical and cyber protection services are also outsourced, making optimized defense more complicated.[33]

Given the trend of emerging physical threats against critical infrastructure, governments and private industry need to be cognizant of the responsibility they shoulder and therefore remain vigilant in their efforts to anticipate and counter physical threats to the greatest extent possible. It is essential that security managers learn to monitor physical threats and constantly innovate for maximum resilience, for the cost of failure is just too high.

Nexus between Threat and Risk

An accurate threat assessment relies on detailed knowledge and understanding of the intentions and capabilities of the potential threat actors, whether an act of nature, a proxy force, malevolent non-state actors, terrorists, or an insider with malicious intent. A quality security risk assessment considers the likelihood—including both threat and vulnerability—and consequences of unwanted events perpetuated by a host of threat actors. Vital to any credible

32. Johnathan Tal, "America's Critical Infrastructure: Threats, Vulnerabilities and Solutions," Security InfoWatch (website), September 20, 2018, https://www.securityinfowatch.com/access-identity/access-control/article/12427447/americas-critical-infrastructure-threats-vulnerabilities-and-solutions.

33. Tal, "America's Critical Infrastructure."

CISR program is the ability to juxtapose a given threat in relation to the risks and vulnerabilities in critical infrastructure systems and the current capabilities of relevant stakeholders. Unless there is strategic commitment to a security program based on threats and risks to critical infrastructure, it will be virtually impossible to create a demonstrably effective national CISR architecture, culture, and policy.

Understanding a threat actor's intentions requires detailed knowledge and understanding of the actor's beliefs, background, worldview, goals, and capabilities to act on any malicious intent by conducting a physical attack. In the case of natural threats or threats due to accidents or technical failures, such knowledge involves scientific predictions of the likelihood of natural disasters or the failure of infrastructure components due to factors like time and stress. Security experts agree the myriad of debilitating and harmful acts that malevolent forces can perpetrate against critical infrastructures are diverse, constantly changing, and becoming more sophisticated and difficult to thwart. However, if critical infrastructure is to be secured effectively and made more resilient, then those stakeholders with this charge need to understand thoroughly the spectrum of plausible physical threats that exist today or lie just over the horizon—a place where the unknown unknowns, introduced in chapter 1, often emerge.

Risk is the potential for an unwanted outcome resulting from an incident, event, or occurrence, as determined by its likelihood—a function of threats and vulnerabilities—and the associated consequences. *Risk management* is the process of identifying, analyzing, and communicating risk while accepting, avoiding, transferring, or controlling it to an acceptable level and cost. Risk management focuses resources on those threats and hazards that are most likely to cause significant, unwanted outcomes to a specific critical infrastructure system or sector, and informs actions designed to prevent or mitigate the effects of those incidents. It also increases security and strengthens resilience by identifying and prioritizing actions to ensure continuity of essential functions and services, and to support enhanced response and restoration. Risk management facilitates decision making and the setting of priorities among all stakeholders.

A risk management framework sets out an approach to fulfill three key functions in a consistent manner. First, the framework aims to identify, analyze, and allocate resources to deter, detect, disrupt, and prepare for threats and hazards to critical infrastructure. Second, the risk management framework enables decisionmakers to prioritize and direct efforts to reduce vulnerabilities, and address physical features

or operational attributes of critical infrastructure systems that are vulnerable to exploitation or susceptible to a given hazard. Finally, it helps stakeholders proactively mitigate the potential consequences of incidents or prepare to limit them effectively if they do occur. The risk management framework can be applicable to all levels of government or private sector organizations. It should cover all threats and hazards and varying factors across critical infrastructure sectors, in addition to individual assets and systems. See chapter 13 for an in-depth explanation of the nature of security risk assessment and management as well as relevant best practices.

The types of physical threats identified in this chapter have the potential to harm, damage, incapacitate, or destroy critical infrastructure. The intent and capability of an adversary, the adversary's access to a critical infrastructure target, and the opportunity to act upon these threat motives contribute to if, when, and how threat groups actually execute an attack. Threats to critical infrastructure vary in intent, capability, potential or intended targets, attack methodologies, and opportunities. Rather than focusing on one type of threat or hazard at a time, nations should identify all threats and hazards that pose the greatest risks to their critical infrastructure. This process allows for more effective and efficient planning and resource allocation.

Critical infrastructure has long been subject to risks associated with many of the physical threats discussed in this chapter, but now serious risks and threats emanate from the cyber domain as well. These risks stem from a growing integration of information and communications technologies with critical infrastructure and from adversaries focused on exploiting potential cyber vulnerabilities. As physical facilities become more reliant on complex cyber systems for operations, critical infrastructure will become increasingly vulnerable to certain cyber threats. Many of these cyber threats can cause similar levels of damage, disruption, and destruction as a number of physical threats. See chapters 3 and 4 for their discussion of the nature of the threats in the cyber domain and the combination of physical and cyber threats in a hybrid manner.

Connections and interdependencies between infrastructure elements and sectors mean that damage, disruption, or destruction to one infrastructure element by a physical (or cyber) event can cause cascading effects that impact continuity of operations in other sectors or systems. Identifying and understanding interdependencies (two-way) or dependencies (one-way) between infrastructure elements and sectors are important for assessing the risks and vulnerabilities, and for determining which steps

may be taken to increase security and resilience. See chapter 12 for further explanation of the dependencies and interdependencies between critical infrastructure sectors, and the potential for cascading or escalating effects. For example, the electric grid relies on integrated information and communication systems from other critical infrastructure sectors in order to operate, while the electric grid also powers lifeline sectors (such as transportation and water treatment systems).

Conclusion

The current threat environment is changing in ways that require new kinds and levels of attention. Critical infrastructure across NATO member states is in both the geopolitical battle space and the target of extensive criminal activities. The foreign intelligence threat has never been more complex or dynamic than it is today. Foreign intelligence entities are developing the capacity to exploit, disrupt, or degrade the critical infrastructure systems that underpin global energy and financial markets, telecommunications services, government functions, and defense capabilities.[34] By these efforts, foreign intelligence services seek to influence or coerce decisionmakers during periods of crisis by holding critical infrastructure at risk. Compared to the rest of the world, NATO member states generally have had the tremendous privilege of being able to take functioning critical infrastructure for granted. Clean water, reliable roads, high-quality health care, dependable electricity, telephones, and e-mail are all so fundamental to modern life in the West it is impossible to picture life without them. However, the events of the first decades of the twenty-first century demonstrate the need to reevaluate any assumptions of uninhibited access to these services. Threat elements will likely plan to avoid Allied armed forces and instead seek to damage those targets which are not well defended and could cause the most physical and psychological damage. Malevolent forces are targeting the infrastructures that support the NATO member states and their armed forces.

For NATO nations, it is vital to understand the nature of these critical infrastructures, their inherent vulnerabilities and risks, and the threats they face. As new technologies increase the speed of operations, the flow of information, or the timeliness of developing a common operating picture, opportunities for threat actors to damage or destroy critical infrastructure also increase. When adopting new systems

34. NCSC, *Insider Threat Mitigation*, 6.

or building new critical infrastructure facilities, Allies and partners must constantly assess the inherent risks and vulnerabilities of these systems as well as the ways in which malevolent forces could threaten them. This vigilant assessment helps ensure that every move forward does not expose a weak link that threat groups can target and attack.

Given the evolving threat environment, critical infrastructure stakeholders in the private and public sectors and all levels government must continually assess their security postures against the spectrum of physical threat actors and threat vectors they face. Assessing the organization's overall enterprise security posture and accounting for recent investments or changes in organizational security are key steps to take. Additionally, organizations must evaluate to what extent their security posture aligns with the current and emerging threat environment. When the security posture and procedures do not align the nature of the threat, it is vital that organizations identify the mismatch and know which leaders in the organization are accountable for correcting such a deficiency in the security program, and how they should go about it. If existing policies and practices cannot answer such questions, critical infrastructure organizations need to conduct a security posture review and assessment.[35] With such a strong commitment to assessing risks and threats, and taking the steps to mitigate them, NATO member states and partner nations can strengthen the culture of security and resilience needed to mitigate the broad array of physical threats to critical infrastructure. Such measures will contribute to the security and prosperity of individual nations and also put the broader Alliance in a better position to fulfill NATO's core tasks.

35. NCSC, *Insider Threat Mitigation*, 10.

— 3 —

Cyber Threats to Critical Infrastructure

Salih Biçakci

The risks against critical infrastructure are on the rise, particularly in the cyber domain. While the COVID-19 pandemic has increased the average household's online activity and compelled businesses to adopt practices to accommodate a more remote workforce, it has also presented malevolent attackers an unprecedented opportunity to test cybersecurity systems and exploit vulnerabilities. Within the first six months of the pandemic, the US Federal Bureau of Investigation's (FBI) Cyber Division reported a 400 percent increase in total cyberattacks compared to pre-pandemic levels, while Interpol highlighted an "alarming rate of cyberattacks aimed at major corporations, governments, and critical infrastructure."[1] The pandemic has also demonstrated the significance of the digital infrastructure for the continuity of modern life. The dependable operation of electricity, natural gas, oil, water and wastewater systems, and telecommunications is vital for the functionality of a state and the essential services it provides. In this sense, critical infrastructure stakeholders are responsible for establishing a cybersecurity posture with implications for national security.

In the cyber domain, two perspectives prevail: an offensive mindset or a defensive mentality. From the defensive perspective, those responsible

1. MonsterCloud, "Top Cyber Security Experts Report: 4,000 Cyber Attacks a Day since COVID-19 Pandemic," CISION PR Newswire (website), August 11, 2020, https://www.prnewswire.com/news-releases/top-cyber-security-experts-report-4-000-cyber-attacks-a-day-since-covid-19-pandemic-301110157.html.

for defending their systems are facing great uncertainty as they seek to counter and mitigate the challenging array of threats in the cyber domain. In a way, the offensive approach is simpler since attackers focus exclusively on finding a system vulnerability to breach and exploit. This chapter prioritizes the offensive perspective and thus will adopt an adversarial perspective to focus on the vulnerabilities, risks, and threats to critical infrastructure in the cyber domain. For an overview of the defensive mentality, see chapter 14 for its discussion of the cybersecurity tools and best practices for securing critical infrastructure.

To design a successful attack, threat actors must comprehend the nature of the target and its relevant components and relationships. From the offensive outlook, the ultimate goal is first to find a single vulnerability to gain access to the computer systems, which the attacker will later exploit to elevate privileges within the system. The offensive mentality reflects security practices abundant in the natural world. In nature, for example, predators use two primary tools—stealth and strategy—to capture their prey.[2] In the cybersecurity sector and associated literature, one helpful depiction of this offensive mindset is Lockheed Martin's Cyber Kill Chain® model, which identifies the seven steps cyber threat actors must take to achieve their goals.[3] Based on its knowledge base of tactics, techniques, and procedures that adversaries have used in cyberattacks, MITRE's ATT&CK® methodology elaborates on these seven steps in a framework that diligently outlines a possible adversarial approach.[4] The matrix outlines the phases of a cyberattack, beginning with an initial period of preparation—consisting of reconnaissance and weaponization—and continuing with key steps such as execution, persistence, privilege escalation, defense evasion, and lateral movement.

Since every critical infrastructure system is situated in a unique operating environment, attackers must first understand the landscape and the targeted organization. Critical infrastructure can be understood as a colossal gear mechanism that only fulfills its specific function if all gears work as designed; any design or organizational flaw, therefore, quickly becomes a vulnerability. Second, most critical infrastructure assets and systems are owned and operated by the private sector. Thus, they are generally business-oriented facilities that must operate within specific economic limitations. Finally, critical infrastructure inherently involves the human

2. See Raphael D. Sagarin and Terence Taylor, eds., *Natural Security: A Darwinian Approach to a Dangerous World* (Berkeley: University of California Press, 2008).

3. "The Cyber Kill Chain," Lockheed Martin (website), n.d., accessed September 2, 2021, https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html.

4. "ATT&CK Matrix for Enterprise," MITRE (website), n.d., accessed September 2, 2021, https://attack.mitre.org.

dimension because critical infrastructure assets and systems are (1) owned and operated by the human workforce, and (2) they provide essential services to end users, so any disruption will impact daily life.

This chapter will provide an overview of a critical infrastructure's technical layers and systems, followed by an examination of its various layers of technology, and then a discussion of its potential organizational vulnerabilities related to the human workforce and management. Next, the chapter will introduce the difference in mentality between cyberattackers and defenders, highlight the various categories of threat actors, and conclude with an overview of the rising and recent primary attacks types to demonstrate the various approaches cyber threat actors employ to exploit the vulnerabilities in critical infrastructure.

Technical Layers and Structures in Critical Infrastructure

Infrastructures are large-scale, manufactured systems that operate interdependently to produce and distribute essential goods—such as energy, water, and data—and services, including transportation, banking, and health care.[5] Critical infrastructure plays an essential role in the vitality of a society and the functionality of the state. See chapter 1 for an overview of the relationship between lifeline sectors and other critical infrastructure sectors. Based on factors like their geography, natural resources, and economies, states differ in which infrastructure systems they classify as critical. See chapter 12 for a comparison of which sectors some countries classify as critical infrastructure. In an oil-producing country, for example, the priority to protect oil refineries is much higher than in others. Each critical infrastructure sector has unique requirements and practices given the different services and goods these sectors provide. Among the various types of critical infrastructure facilities, there are differences in design, planning, functionality, organization, and specification properties depending on the sector. For example, the design and function of nuclear power plants are notably different from telecommunications or water systems. For a better grasp of cybersecurity of critical infrastructures, it is important to note that—despite the differences outlined above—systems also have similarities on the organizational and technical layers that merit further exploration.

5. Enrico Zio, "Critical Infrastructures Vulnerability and Risk Analysis," *European Journal for Security Research* 1 (2016): 99, https://link.springer.com/content/pdf/10.1007/s41125-016-0004-2.pdf.

Connectedness and Technical Complexity

By nature, critical infrastructures are inherently complex, distributed systems that resemble a tree with its roots buried in the soil and its branches in plain sight. It is therefore more appropriate to model the interconnectedness of these structures as a system-of-systems, which consists of "multiple, heterogeneous, operationally distributed, occasionally independently operating systems embedded in networks at multiple levels that evolve over time."[6] The complexity of a system-of-systems affects both the protection requirements and crisis management procedures. There are two significant types of complexity present in these critical infrastructure systems. First, the level of connectedness between systems means that any malfunction could create a cascading effect among similar facilities. Second, the connectedness also affects other sectors that are not similar in kind but nonetheless dependent. See chapter 12 for a detailed discussion of cascading and escalating failures resulting from dependencies and interdependencies between critical infrastructure systems. For instance, any glitch in the telecommunication sector would quickly affect all sectors that depend on telecommunications for their operations. To illustrate this concept of connectedness, figure 3-1 shows how a power outage in Italy in 2003 affected several other critical infrastructure sectors within Italy and across national borders.[7]

6. Daniel DeLaurentis, "Role of Humans in Complexity of a System-of-Systems," in *Digital Human Modeling*, ed. Vincent G. Duffy (Berlin: Springer, 2007), 363.

7. Wolfgang Kröger and Enrico Zio, *Vulnerable Systems* (London: Springer, 2011), 13.

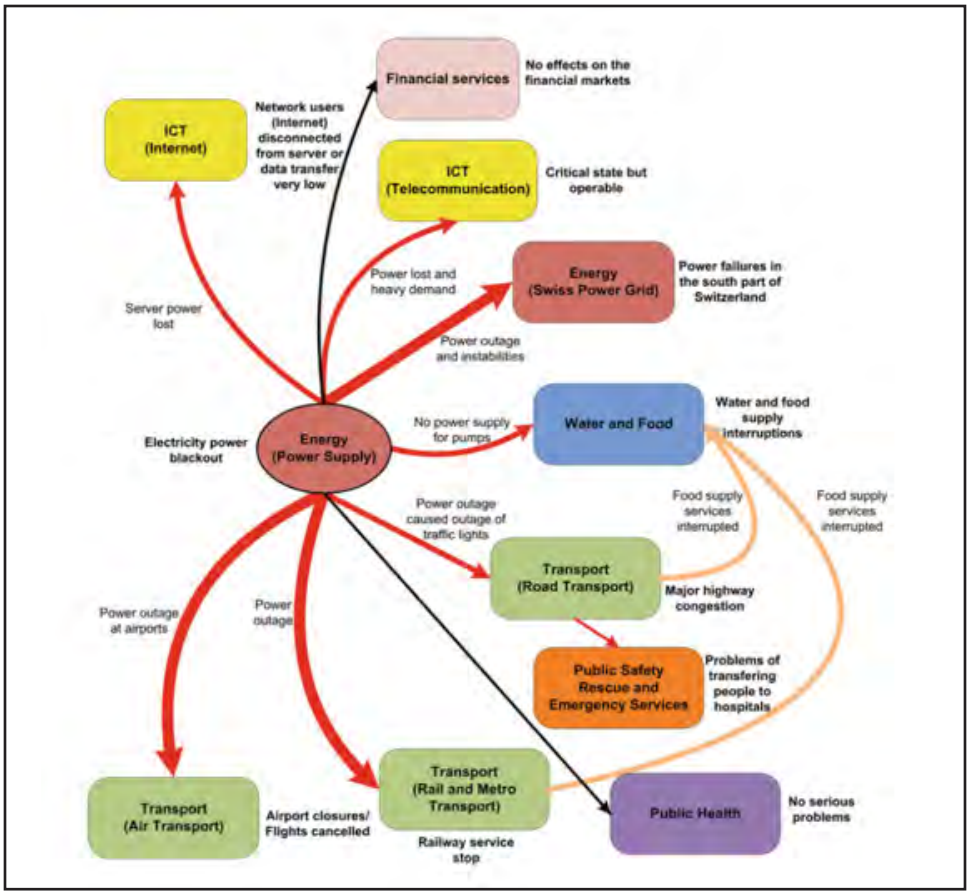


Figure 3-1. Impacts of the Italian blackout on other sectors
(Diagram by Springer)

The complex nature of these dependencies, which increases the potential for significant crises and cascading or escalating failures, highlights the need for cross-sector coordination, planning, and exercises before cyberattacks occur.[8] Unfortunately, the general lack of cooperation, communication, and participation in comprehensive exercises across critical infrastructure sectors is a vulnerability that malevolent cyber actors could exploit. One security expert notes:

> Complex systems like the electric power grid, water networks, transportation networks, and communications networks tend to self-organize into a critical state, and, once in this state, any

8. Gianluca Pescaroli and David Alexander, "Critical Infrastructure, Panarchies and the Vulnerability Paths of Cascading Disasters," *Natural Hazards* 82 (2016): 187–88, https://link.springer.com/content/pdf/10.1007/s11069-016-2186-3.pdf.

> change to the system can start a chain reaction. These chain reactions manifest as cascading avalanches, nuclear power plant meltdowns, and electrical power grid collapses. When a sand pile reaches a critical state, the addition of a single grain of sand may lead to avalanches of unpredictable size—even extreme avalanches that completely destroy the sand pile.[9]

Communications between and among government authorities and critical infrastructure owners and operators is, therefore, a crucial element of CISR policy and practice. See chapter 11 for examples of multidirectional information sharing and blue-sky coordination between the public and private sectors. In a critical infrastructure facility, there are several unique levels of operation, including business, management, core production, supply chain (third parties), and perimeter security. To sustain the proper functionality of the facilities, coordination and communication between these layers are critical. In the core production layer, two types of technologies are essential: information technologies (IT) and operational technologies (OT). Although similar in some ways, IT and OT are each unique enough to require operators and defenders to have a particular knowledge and special insight to work with them. The integration of these technologies is a vital element of robust cyber systems.

In addition to the operational complexity of connectedness, there is a temporal effect on these facilities. The industrial control system (ICS) architecture is comprised of layers of solutions, processes, and procedures. Underlying and ensuring the sustained operations of critical infrastructure is an ICS structure that consists of several types of control systems, such as supervisory control and data acquisition (SCADA), distributed control systems, and programmable logic controllers. A hallmark of Industry 4.0 is the prevalence of Internet of Things (IoT) components in the critical infrastructure domain. The increase of wireless communication between nonhuman components also pushes the limits of wireless communication protocols and creates additional responsibilities in cybersecurity.

In ICS management, several components have to function together to sustain the operation of critical infrastructure. Adding to the complexity of the ICS, the critical infrastructure network is comprised of several IT and OT components made by different vendors and typically with different lifespans, dissimilar updates, and varied support requirements. When critical infrastructure owners and operators prioritize

9. Ted G. Lewis, *Critical Infrastructure Protection in Homeland Security* (Hoboken, NJ: Wiley, 2015), 49.

functionality and productivity over the security of ICS components, cybersecurity risks increase in the critical infrastructure ecosphere. Thus, when it comes to implementing CISR policy and ensuring continued operation of critical infrastructure, the facility's chief security officer plays a crucial role in mitigating the vulnerabilities, risks, and threats in the cyber domain. ICS components are not publicly circulated devices or components, so targeting them requires attackers to have intimate knowledge of the components—a reality that decreases the number of adversaries with the degree of specificity needed to attack OT components. On the other hand, the attacks targeting ICS components have predominantly originated from advanced persistent threat groups or those with state sponsorship.

The ICS structure, built by components from different vendors, requires considerable planning to integrate the necessary updates, upgrades, and security patches and to ensure this renewal process does not hinder the overall function of the system. Since it can be challenging to know how these various components will function after completing the renewal process, those responsible for these upgrades and updates could be hesitant to introduce any change that may threaten the system's ability to work as designed. As a result, the SCADA and ICS components that enable the operation of critical infrastructure often do not receive most of the updates and security patches they require.

Layers of Technology: Information Technology, Operational Technology, and the Industrial Internet of Things

The functionality of different computer protocols within a given critical infrastructure enhances the system productivity but complicates the security. The IT and OT systems within a critical infrastructure's core production area are connected and synchronized to sustain the production or servicing of the facility. As its name indicates, designers of OT have in mind the ability to perform specified actions that ensure operational continuity of the critical infrastructure. As a result, during the operational design of OT systems, functionality is the priority, not security. Engineers, in principle, do not connect the OT systems directly to the Internet during the design phase, but use firewalls to keep OT systems separate. Since information technology focuses on the computer and telecommunication systems that perform data input, storage, recovery, transmission, data processing, and data protection, it prioritizes data and the proper data handling procedures. Innovations in the IT realm change and improve rapidly, and most IT infrastructures could quickly embrace and adopt the latest technological advancements with little effort. In a critical infrastructure

environment, an IT network forms the foundation for the business and enterprise levels while the tightly coupled nature of IT and OT systems in the core production area presents an important vulnerability.

In the 1980s, OT and IT systems functioned primarily in an independent manner. In the 1990s, OT systems started to connect to IT systems to centralize the management. In the 2000s, OT engineers built IT-compatible systems to improve the communication problems between two systems. The boundaries between IT and OT systems thus started to blur in the networks. Today, OT systems also cooperate with cloud computing and wireless technologies.

Simply put, operational technology is the computing capacity to fulfill an operation and monitor devices, various industrial processes, and some industry events. See figure 3-2 for a typical network architecture for OT industrial systems.[10] OT systems perform important functions, including production-line management, mining-operations control, and oil and gas monitoring. The operational technology prioritizes the process required to keep the system functional, so enhancing the availability, precision, and reliability of OT services is essential. In the critical infrastructure sector, OT systems regularly monitor the processes and support the manufacturing and defense utilities.

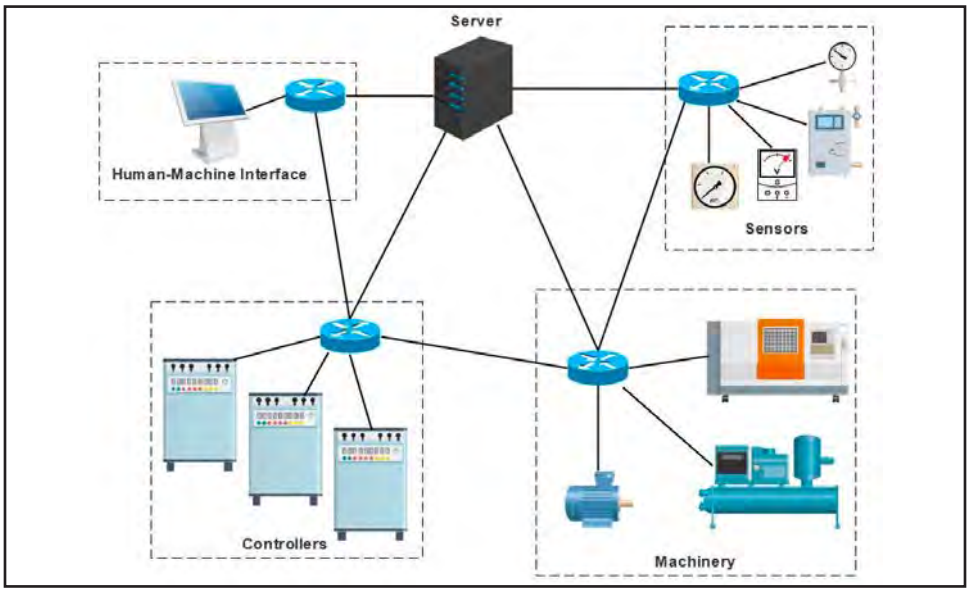


Figure 3-2. Typical OT industrial systems network architecture
(Diagram by Applied Sciences)

10. Jaco Prinsloo, Saurabh Sinha, and Basie von Solms, "Review of Industry 4.0 Manufacturing Process Security Risks," *Applied Sciences* 9, no. 23 (2019): 3, https://www.mdpi.com/2076-3417/9/23/5105/pdf?version=1574739445.

Unlike most IT systems, which are highly controlled and regulated by system administrators, OT systems typically are not governed by the same limitations or access rules. To protect the functionality of OT systems, they should not be as accessible as the IT systems. Compared to IT systems, OT systems are far more difficult to recover and restore in the event of a failure, and any OT network failure can yield catastrophic results or lead to an "end-of-life" scenario.

Over time, OT systems have gradually become more secure, though this progress does not match the considerable efforts focused on securing IT systems. With limited firmware upgrades, the supported lifetime of an OT component is between 10–20 years. In contrast, the hardware, software and protocols of IT systems are rapidly changing. This difference in the pace and types of changes between these two systems creates compatibility and communication problems. The technological changes in IT and OT systems can challenge critical infrastructure operators' ability to adapt to the changes and manage the process of installing updates and patches. Given these challenges, critical infrastructure facilities require strategic plans to manage this process and ensure the proper blend of security upgrades and continued business operations.

A significant segment of OT systems is the ICS architecture, which consists of systems that monitor and control industrial processes with the assistance of programmable logic controllers or discrete process control systems. SCADA systems, which are responsible for managing the ICS, have two main components. The first is the human machine interface. The second is the historian, which provides a graphical user interface for operators to observe the status of a system easily, receive any alarms indicating out-of-band operation, or enter system adjustments to manage the process under control. In addition to SCADA, key programs in the OT systems category include computer numerical control, building-management systems, and building-automation systems. The OT systems also have their own communication protocols that differ from those used in IT systems. OT systems also frequently use remote terminal units (RTU) to connect one or multiple devices (actuators and monitors) to manage or monitor a process from close proximity or from thousands of kilometers away from the headquarters.

Beyond IT and OT systems, the newest element influencing the operations and cybersecurity of critical infrastructure is Industry 4.0, which enables the intelligent networking of machines and processes to increase efficiency. The Industrial Internet of Things (IIoT) refers to the

use of the IoT in industrial sectors, and it enables machine-to-machine and artificial-intelligence applications in several categories. Although not without risks, these machine-to-machine applications can simplify the complicated production or control processes. In critical infrastructures, the primary distinction between IIoT devices lies at the intersection of IT and OT layers. The IIoT can also optimize visibility of the supply chain and logistics by utilizing smart sensors and actuators. This networking structure can also enhance the OT security community's knowledge and understanding of its networks.[11] In the OT environment, the challenge is determining whether the problem is an unusual incident resulting from an attack or simply a basic software error.

Social Complexity and Socio-technical Structures

While the technical complexity of computer systems has been a topic of study for years, the concept of social complexity is unique and comparatively new. A conventional understanding of security focuses on securing nonhuman components in the operational zone and looks to automation as the ideal tool to control all possible outcomes. The automation and restriction of nonhuman components require less effort than dealing with social unpredictability. The problem is that critical infrastructures are business-oriented service sectors and have to cooperate with several third-party partners. It is possible to leverage automation as a means to minimize the human workforce, but not without exposing the critical infrastructure to other vulnerabilities.[12] The modern approach to social complexity, however, understands the whole of the network and its human and automated components within the facility. Engineers, together with their cell phones, access cards, tokens, and server rooms, should be considered in a networked mentality. This perspective would reveal just how many components are connected and demonstrate the levels of complexity in a facility.

The actor-network theory (ANT) understands each element as a node in a particular organization or structure.[13] The micro-level understanding

11. David Masson, "SANS ICS Security Summit 2021 Recap: Industry on the Move," *Darktrace* (blog), March 26, 2021, https://www.darktrace.com/en/blog/sans-ics-security-summit-2021-recap-industry-on-the-move/.

12. Federico Maggi and Marcello Pogliani, *Rouge Automation: Vulnerable and Malicious Code in Industrial Programming* (Irving, TX: Trend Micro Research, 2020), https://documents.trendmicro.com/assets/white_papers/wp-rogue-automation-vulnerable-and-malicious-code-in-industrial-programming.pdf.

13. Karin Hedström, Gurpreet Dhillon, and Fredrik Karlsson, "Using Actor Network Theory to Understand Information Security Management," in *Security and Privacy—Silver Linings in the Cloud*, ed. K. Rannenberg, V. Varadharajan, and C. Weber (Berlin: Springer, 2010), 43–54, https://link.springer.com/content/pdf/10.1007/978-3-642-15257-3_5.pdf.

of the critical infrastructure in a networked mode can help the protection level see and assess the vulnerabilities and weaknesses more easily by giving clues about the specific interactions between the various elements in a facility. The ANT approach also can present a holistic perspective on the security of critical infrastructure. When an attacker decides to target a specific critical infrastructure, the next step will be to find the right vulnerability to exploit and thus gain initial access to the system. Attackers pursue this goal via hardware additions, trusted relationships, replication through removable media, and external remote services.[14] In this context, the theft of an engineer's mobile device or unauthorized access to a manager's virtual private network accounts could quickly become vulnerabilities that jeopardize the functionality of the entire system.[15] The practice of ANT can also assist critical infrastructure operators and cybersecurity teams to identify potential consequences if and when they happen.

Additionally, temporal changes can also impact critical infrastructure systems. There are times when economic and political considerations or natural disasters compel critical infrastructure operators to expand the facility's output capacity to respond to new service demands for a period of time. There are also situations in which a facility begins under public ownership but then transfers to private sector management or when the layout of the facility changes for reasons such as renovations or additional construction. High turnover among the workforce at these facilities can affect the requisite expert knowledge to ensure efficient operations, even if specific blueprints and records are available to new workers.

In most critical infrastructures, the nature of operations—namely, uninterrupted service—means there is reduced visibility and control during the system's functionality. Facilities tend to coordinate their production levels with the facilities in kind to organize the distribution of the product that customers demand. Electricity, watering, sewage, and telecommunications sectors primarily work on similar principles. From an organizational structure, critical infrastructure facilities are mainly divided into two sections: (1) the production plant and the management layer, and (2) the services and distribution lines. For the first category, various equipment

14. Sergey Golovanov, "DarkVishnya: Banks Attacked through Direct Connection to Local Network," SecureList (website), December 6, 2018, https://securelist.com/darkvishnya/89169/; and "Steal macOS Files with the USB Rubber Ducky," Null Byte (website), July 14, 2017, https://null-byte.wonderhowto.com/how-to/steal-macos-files-with-usb-rubber-ducky-0177336/.

15. Kensington, *Locking Down Mobile Devices to Keep Business Data Safe: Physical Security Solutions Are Bridging the Gap between Productivity and Protection* (San Mateo, CA: Kensington, 2018), https://www.kensington.com/siteassets/documents/kensington-lockingWP-277450-june2018-FINAL.pdf.

and tools limit physical access to the facility where production occurs. Inside the plant, the physical and cybersecurity teams work to control any anomaly to prevent breaches while operators monitor sensors to manage the operation of the facility. In the second category, the distribution lines and services have limited protection from any threat because the attack surface is significant due to its vast size. In some sectors, security of these different lines and services relies on something as minor as a lock pad or a protected box to deter a would-be adversary from attacking or targeting the system. An expert could easily utilize these nodes to harm or disrupt the critical infrastructure and the services it provides.

For example, many electricity and telecommunications companies use service cars and trucks that typically remain parked in the facility garage under limited perimeter security or observation during off hours. From a business perspective, these companies seek cost-effective means to work and store equipment. This minimal level of security, however, offers threat actors an opportunity to conduct hostile reconnaissance or obtain identification badges, working vests, helmets, and even work computers or special tools in some cases. An inherent vulnerability of facilities with such a comprehensive network is the difficulty in protecting all personnel, locations, and equipment at all times. The expansion of the network that enlarges the attack surface is also a vital consideration for CISR planning and efforts.

Seeking Gaps in the Organization and Business Management Levels

Human Capital, Culture, and Security

In critical infrastructures, the human element is one of the most significant factors to the system's operations and effective functionality. Initial steps in planning and building critical infrastructure generally prioritize details related to perimeter security, digital security, and operational safety with comparatively less emphasis on the necessary investment in the human element. Most of the key security concepts—such as situational awareness, strategic communications, suspicion, resilience, and a culture of security—are fundamentally rooted in the critical infrastructure's human capital: the people who own, manage, and operate the system. Since even the most sophisticated systems are worthless without a qualified human workforce,

some researchers define the human component as one of the interdependencies of critical infrastructure.[16]

Understanding the human dimensions of the workforce, both on duty and off duty, is essential to an effective CISR practice. In most critical infrastructure facilities, there are three major categories for the workforce: technical, nontechnical, and management. In each category, these subordinate teams leverage their unique skills and capabilities, and cooperate to achieve a common goal. In this way, each team develops its own working culture, while the entire organization also builds a security culture that aims to mitigate threats and build resilience. The human resources (HR) department is the unit primarily responsible for understanding and organizing the workforce. In practice, however, HR departments focus mainly on background security checks rather than understanding individual differences in the workforce that can contribute to or hinder cooperation, which are key elements to the overall security posture. One vulnerability inherent to critical infrastructure is the lack of situational awareness, which is "the knowledge of where you are, where other friendly elements are, and the status, state, and location of the enemy."[17] Within this definition are three levels of situational awareness: (1) perceiving the critical factors in the environment, (2) understanding at those factors mean, particularly when integrated with the decisionmaker's goals, and (3) anticipating how these factors will influence the system in the near term.

Building cooperation among the elements of the workforce requires specific coordination and communication skills that form the pillars of situational awareness. Often, the HR department is best suited to undertake these efforts to improve the organizational culture, communications, and cooperation. Without intentional effort to improve cooperation between departments and their personnel, it is possible for gaps or blind spots to form within the organization, leading to vulnerabilities that adversaries could easily exploit.[18] Efforts to build normal, healthy communications and cooperation between departments can prevent the formation of departmental subcultures at the expense of organizational cultures and enhance the overall CISR posture within the organization.

16. Joshua Barnes and Kenneth Newbold, "Humans as a Critical Infrastructure: Public-Private Partnerships Essential to Resiliency and Response," First IEEE International Workshop on Critical Infrastructure Protection (Piscataway, NJ: IEEE, 2005), https://doi.org/10.1109/iwcip.2005.13.

17. Brian T. Bennett, *Understanding, Assessing, and Responding to Terrorism: Protecting Critical Infrastructure and Personnel* (Hoboken, NJ: Wiley & Sons, 2007), 292.

18. Ross Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems* (Indianapolis: Wiley & Sons, 2020), 79.

Most organizations have red teams that serve to test an organization's security posture and pursue ways to enter a facility by physical or cyber methods. In a realistic example, red team members could use a company vest, hard hat, or fake identification badge and approach the entrance as a subcontractor or third-party services maintenance team. In this attempt, the red team might use valid and accurate data obtained with various methods to convince the gatekeeper to authorize entry.[19] Any lack of coordination among the departments—such as notifying security of a lost identification card or stolen uniform items—could enable a true malevolent actor, not just a red team member, to access the facility. To build a robust security culture and framework, HR departments can help foster trust among the workforce, enable open communications, and increase coordination. Empowering individual workers and building a strong organizational culture promotes personal responsibility and encourages the workforce to take ownership of conditions in the facility, including security. See the discussion of airport security communities in chapter 6 for a similar concept. Failing to address these differences between departments can lead to vulnerabilities and failures that red teams or true threat actors can exploit with social engineering techniques.

Ultimately, the management level is responsible for the organizational culture and security at the facility. Owners and managers should take steps to improve trust and cooperation between departments and also encourage a culture in which the workforce can contribute to and, when appropriate, participate in the decision-making process. These efforts can engage the organization's human capital and encourage the workforce to report problems when they arise, especially from the security perspective. This participatory approach would also enable change management and ease the process of adaptation required to enhance CISR policies and practices.

Business Management and Coordination in Critical Infrastructure

The business layer is one of the central departments in critical infrastructures. Since many sectors of critical infrastructure are under private ownership—and thus earn a profit from their services—the business mindset sometimes prevails over core security and safety practices. In addition to coordination and communications within the organization, the management level is also responsible for fostering secure and healthy communications with the external regulatory bodies to ensure the facility's continued operations. These communications also include measures to share information

19. Wil Allsopp, *Unauthorised Access: Physical Penetration Testing For IT Security Teams* (West Sussex, UK: Wiley, 2009), 29.

and intelligence regarding possible threats. See chapter 11 for a discussion of multidirectional information and intelligence sharing. Figure 3-3 depicts the role and key relationships the business layer has within and outside the organization.

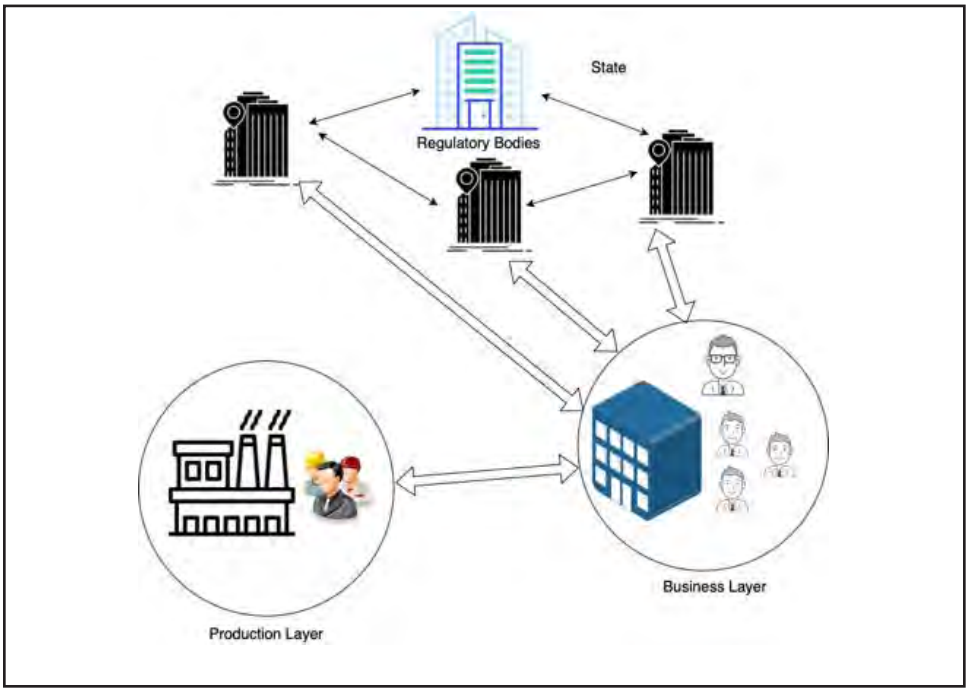


Figure 3-3. Role of the business layer

The business layer is also responsible for managing the facility's effective operations in the face of threats such as cyberattacks. One of the most rewarding investments managers can make is to conduct regular, unannounced exercises to train employees, enhance cooperation, and minimize the effects of a cyberattack. Such exercises expose the workforce to some of the uncertainty and stress of a real cyberattack and instill proper responses, improve decision making, and make these behaviors a more natural reflex. Exercises can also expose cognitive biases in the workforce, including overconfidence in friendly capabilities and an underestimation of threats to the organization's systems. Unfortunately, the turnover in the workforce, the demands of daily business, and the perception that exercises are too disruptive combine to limit the conduct of these valuable exercises and the benefits they can yield. If the business layer blindly trusts the technological investments and ignores the human component in these socio-technological systems, then the organization is more susceptible to damaging cyberattacks.

Finally, the business layer is the focal point for leading change management initiatives. While the business mentality resists making changes that may disrupt the facility's operations, updating the digital components is essential to cybersecurity in critical infrastructure. In fact, evidence indicates that implementing necessary changes generally offers substantial benefits to critical infrastructures and outweighs the costs for these upgrades.[20] To help senior managers initiate and sustain significant changes, it is vital to define the roles and responsibilities for the various layers of the organization and the workforce across different departments.[21] Executive leaders should also gain a better understanding of the technical layer to manage the facility during crises and mitigate likely threats.

Mindsets and Threat Actors

A Difference in Mentality: Attackers and Defenders

Successful security efforts rest on three key pillars: (1) precautions to guard and protect the design and structure, (2) sensors and alarms to alert the systems in the event of a breach or any malign activity, and (3) human capacity. Even if a critical infrastructure is flawless in its design and able to detect any abnormality, only the human dimension—managers, operators, and the broader workforce—can determine the appropriate responses when problems arise. The threats malicious cyber actors pose to critical infrastructure are complex, as this analogy describes:

> Today's cyber security environment is like playing 1,000 simultaneous chess matches against different opponents of varying skill levels. While it is simple to defend the board against 98% of the more junior players (script kiddies), the top 2% require real effort and strategy.[22]

In the security of critical infrastructure, offensive and defensive mindsets define the rules in the cyber domain. Defenders do not know who

20. Thomas Lauer, *Change Management: Fundamentals and Success Factors* (Berlin: Springer, 2021), 69, https://doi.org/10.1007/978-3-662-62187-5.

21. Software Engineering Institute, *Configuration and Change Management* (Pittsburgh, PA: Carnegie Mellon University, 2016), 10, https://us-cert.cisa.gov/sites/default/files/c3vp/crr_resources_guides/CRR_Resource_Guide-CCM.pdf.

22. Stephen Horvath, "Effective Cyber Defense Is More about Mindset than Budget," Telos (website), March 19, 2015, https://www.telos.com/2015/03/effective-cyber-defense-is-more-about-mindset-than-budget/.

will attack the system, when they will do so, or how they will go about it. Since the threat vector is not readily apparent, defenders must maintain a sharp focus 24 hours per day in all relevant fields of protection, and any signals have to be handled with particular caution due to the uncertainty of the threat. Defenders conduct highly repetitive work—without knowing whether or not their efforts are yielding positive results—that can become so routine that defenders are distracted and lose focus of signs of adversarial activity. The mission of cyber defense does not provide immediate gratification, nor is it well understood by critical infrastructure operators and managers. On paper, it is easy to sustain such an effort, but in reality, cyber defense over an extended period of time is a challenging mission. On the other hand, attackers have a goal in mind, the motivation to succeed, and time to test all possible ways to achieve their desired outcome. Throughout the targeting and attack process, adversaries are engaged, focused, and attentive to details that enable them to find and exploit system vulnerabilities. Attackers are typically not prone to the same temptations of dullness or routine that defenders face; rather, the offensive mentality is marked by a strong motivation to overcome obstacles.

Although training and exercises have an important place in developing security and resilience in critical infrastructure, genuinely simulating potential attackers is not an easy task. As a sort of "devil's advocate," red teams play a crucial role in demonstrating the possible consequences of adversarial attacks. Red teams' cyber and physical penetration tests primarily target the management level of the facility and focus predominantly on issues that could lead to a disruption of normal business operations. Often, red teams inform relevant parties prior to any simulated attacks, and they are bound by specific ethical rules in the conduct of the attack. Here, being ethical and demonstrating a genuine attack psychology are mutually exclusive because real-world attackers have no ethical code to consider or obey. This duality of the security mindset cannot easily be changed with conventional tools. In the age of hybridity and asymmetry, attackers hold several advantages over defenders. One crucial advantage for attackers is the vertical hierarchy of public sector or state-run critical infrastructure systems. When faced with an abnormality or attack, defenders in the targeted entity must inform several managerial groups, and the facility's senior management must notify the state authorities to deal with the consequences of such an attack. Any attack, however, requires a time-sensitive response and prompt action to ensure the facility continues its service.

Threat Actors

For those being attacked, a fundamental and natural question is: Who is the attacker? Demonstrating a conventional and traditional understanding of cyber threats, the cybersecurity literature outlines specific categories of attackers and identifies hackers, crackers, lamers, script kiddies, and lone wolves as the main threat actors. Although the actors along this spectrum vary in their capacities, these distinct categories are no longer relevant because today "the question of knowing the full name of the attacker becomes less relevant than knowing who the enemy is and who the sponsors are; for instance, a state actor or a terrorist organization."[23]

Given their complex networks, unique organizational structures, and role in sustaining modern life, critical infrastructures are distinct from other types of targets. Researchers indicate a rising trend of low-sophistication attacks against critical infrastructure—particularly ICS and Internet-exposed OT—carried out by amateur threat actors who lack expertise and deep knowledge in the sector.[24] In one such example, attackers claimed to have successfully penetrated a German rail-control system, but in they reality had only compromised a web interface for a model train set.[25] Despite such failures by amateurs, a broad range of threat actors seek to attack critical infrastructure, each with different capability levels and motivations.[26] A general profile of potential attackers includes the following groups.

- Opportunistic attackers. These attackers have no specific target but try to spread their malware as much as possible to increase the chance of success.

23. Clement Guitton, *Inside the Enemy's Computer Identifying Cyber Attackers* (New York: Oxford University Press, 2017), 3.

24. Keith Lunden, Daniel Kapellmann Zafra, and Nathan Brubaker, "Crimes of Opportunity: Increasing Frequency of Low Sophistication Operational Technology Compromises," Mandiant (website), May 25, 2021, https://www.mandiant.com/resources/increasing-low-sophistication-operational-technology-compromises.

25. Dawn Blizard, "Critical Infrastructure Attack Trends: What Business Leaders Should Know," Security Intelligence (website), August 19, 2021, https://securityintelligence.com/articles/critical-infrastructure-attack-trends-business-leaders/.

26. "5 Cyber Attack Motives Your Industry May Face," *Otorio* (blog), July 10, 2019, https://www.otorio.com/blog/5-types-of-cyber-attackers-your-industry-may-face/.

- Industrial opportunistic attackers. A variation of opportunistic attackers who target the industrial sector in particular. These attackers try to exploit any zero-day attack or announced vulnerability to catch their targets unprepared with relatively little effort.[27]
- Competitors. These attackers target and steal infrastructure data, such as blueprints or technical information. These events often involve international competitors from within the sector or nation-state intelligence services that practice cyber espionage, and it is rather difficult to distinguish between these two parties.
- Insider threats. These attackers are current or former disgruntled employees who make their attack for a variety of reasons.[28]
- Advanced persistent threats. These attack groups can arrange large-scale, advanced attacks and are tracked by the security community, but are difficult to stop.[29] These highly qualified attackers could be state-sponsored actors intending to execute cyber sabotage or cause chaos in the target country by disrupting or stopping the service.
- Hacktivists. These attackers have ideological goals in mind that legitimize their attack. They are not interested in the results, but in furthering their ideological position by undermining the target organization.[30]

Critical infrastructure owners and managers can also use a simple scale to assess the possible levels of cyber threat against the facility—from one (unsophisticated) to five (advanced), for example—to determine the appropriate threat level against which they plan to defend.[31] Determining this appropriate level of defense allows critical infrastructure management teams to calibrate

27. Bob Rudis, "The Dynamic Opportunistic Attacker Landscape," OPTIV (website), March 17, 2020, https://www.optiv.com/insights/discover/blog/dynamic-opportunistic-attacker-landscape.

28. Marco Rocchetto and Nils Ole Tippenhauer, "On Attacker Models and Profiles for Cyber-Physical Systems," in *Computer Security—ESORICS 2016*, ed. Ioannis Askoxylakis et al. (Cham, CH: Springer, 2016), 427–49.

29. "Groups," MITRE ATT&CK (website), n.d., accessed on September 5, 2021, https://attack.mitre.org/groups/.

30. Sara Ligaard, Norgaard Hald, and Jens Myrup Pedersen, "The Threat of Digital Hacker Sabotage to Critical Infrastructures," in *Image Processing and Communications Challenges 5*, ed. Ryszard S. Choras (Cham, CH: Springer, 2014), 379–90.

31. Deb Bodeau, Jenn Fabius-Greene, and Rich Graubart, "How Do You Assess Your Organization's Cyber Threat Level?," MITRE (website), accessed on August 22, 2021, https://www.mitre.org/sites/default/files/pdf/10_2914.pdf.

their cybersecurity measures to the match the level of threats they face. Ultimately, this step gives managers a better understanding of their capacity to defend against cyberattacks, helps them make necessary preparations and investments, and directs their responses if attacks exceed the projected threat levels.

Current and Emerging Cyber Threats

Every year brings new attack vectors or modus operandi for the threats that mark the era. There are also common methodologies that threat actors prefer to practice that are becoming increasingly popular. This reality means more attack tools are readily available to more potential attackers through an easy and quick online search. In addition to primary cybersecurity essentials, it is crucial to understand the threat landscape to protect networks and enhance CISR posture. IBM's 2021 Security X-Force report, which depicts the most prevalent types of cyberattacks in 2019–20, demonstrates an increase in the use of ransomware, remote access Trojans, and business e-mail compromise in the past few years (see figure 3-4).[32] Additionally, the escalation in data theft, which is also associated with phishing and social engineering, is a methodology attackers prefer, especially in preparation for more sophisticated attacks. Among these various attack types, this section will examine the use of ransomware, business e-mail compromise, credential stuffing, and supply-chain attacks against critical infrastructure targets.

32. IBM Security, *X-Force Threat Intelligence Index* (Armonk, NY: IBM Corporation, 2021), 7, https://www.ibm.com/downloads/cas/M1X3B7QG.

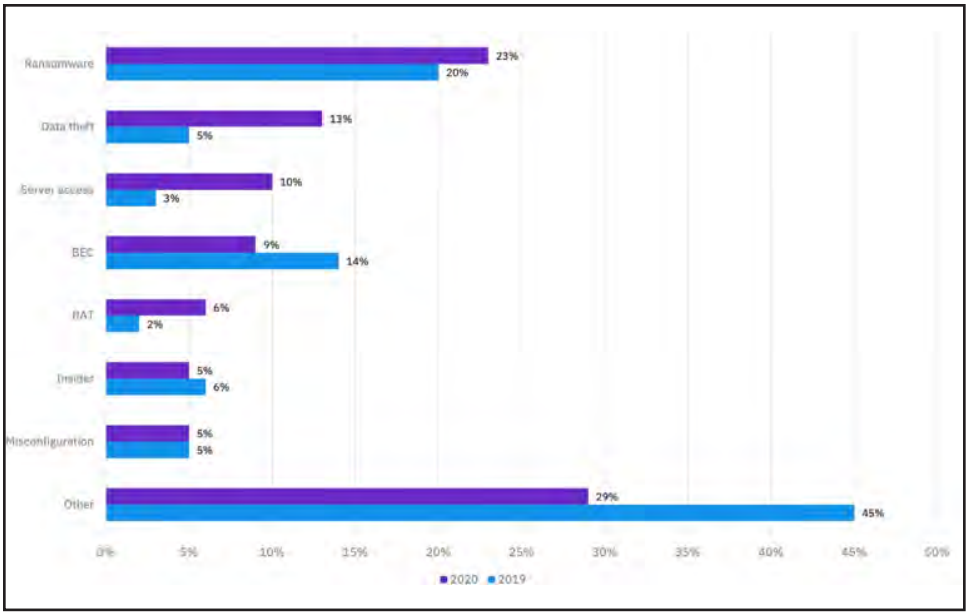


Figure 3-4. Breakdown of cyberattacks by type, 2019–20
(Diagram by IBM)

Ransomware

Ransomware is not simply a malign software that targets a user's data, it is a combination of two concepts: ransom and ware. First, *ware* refers to the malware or malicious software that allows ransomware to encrypt all the data in the infected computer system with various levels of asymmetric encryption. Second, *ransom* alludes to ransomware's targeted campaign and particular business plan to hold a target's data captive and release it only after receiving a payment. Research indicates that private sector businesses experienced a ransomware attack every 11 seconds in 2021, up from one every 40 seconds in 2016.[33]

Ransomware types of software design have been on the market since 1989. That year, Joseph L. Popp conducted the first-ever ransomware attack when he distributed 20,000 floppy disks alleged to contain information on AIDS to the attendees of the World Health Organization's International AIDS Conference. The Trojan encrypted the names of the files on the customer's computer and hid the directories and then demanded $189 to give owners access to their files. The phases of ransomware attacks, like encryption and asking for a ransom payment to restore access, have essentially remained the same

33. "2021 Must-Know Cyber Attack Statistics and Trends," *Embroker* (blog), n.d., accessed on September 29, 2021, https://www.embroker.com/blog/cyber-attack-statistics/.

over time, though the distribution methodology has changed. Today, ransomware has a high likelihood of success with a relatively low chance of discovery. As a result, ransomware has become a tool of choice for cyber threat actors to make money with limited efforts. The increase in ransomware campaigns against critical infrastructure has risen since 2019 and has accelerated during the COVID-19 pandemic (see figure 3-5).[34]

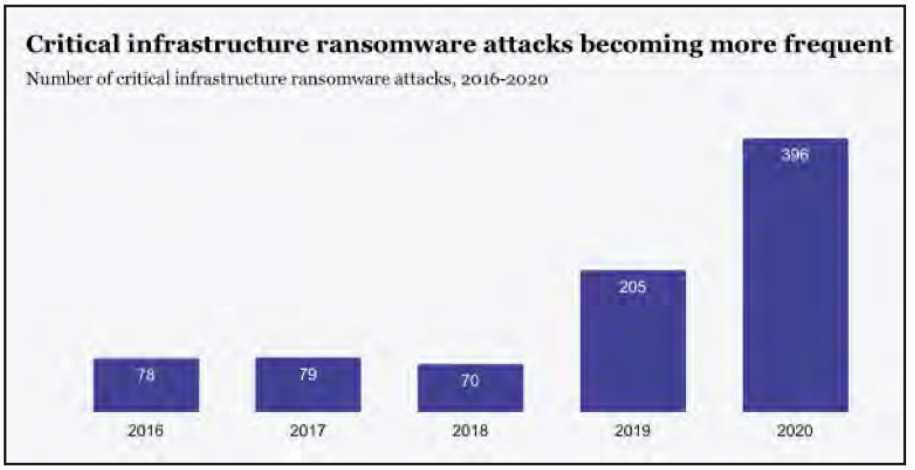


Figure 3-5. Rise of ransomware attacks against critical infrastructure
(Diagram by Temple University)

A recent example of a major ransomware attack occurred on May 7, 2021, against Colonial Pipeline, one of the leading American pipeline operators and sources of fuel for the eastern United States. The attack halted the pipeline's operations, caused fuel shortages and panic buying at gas stations across the region, and led to delays in scheduled airline flights. The Colonial Pipeline attack represents the first time a cyberattack using ransomware affected so many people in their normal daily routines.[35] The success of the Colonial Pipeline attack and the recent upward trend indicate that ransomware attacks are here to stay. So, it is important to understand the nature of a ransomware attack to enhance CISR posture and prepare for these attacks. Figure 3-6 provides a helpful depiction of the life cycle of a ransomware attack and the broad stages it encompasses.[36]

34. Aunshul Rege, "Critical Infrastructure Ransomware Incident Dataset Version 11.4," Temple University (website), n.d., accessed on September 5, 2021, https://sites.temple.edu/care/ci-rw-attacks/.

35. Charlie Osborne, "Colonial Pipeline Attack: Everything You Need to Know," ZDNET (website), May 13, 2021, https://www.zdnet.com/article/colonial-pipeline-ransomware-attack-everything-you-need-to-know/.

36. "How Ransomware Happens and How to Stop It," CERT NZ (website), n.d., accessed on August 23, 2021, https://www.cert.govt.nz/it-specialists/guides/how-ransomware-happens-and-how-to-stop-it/.

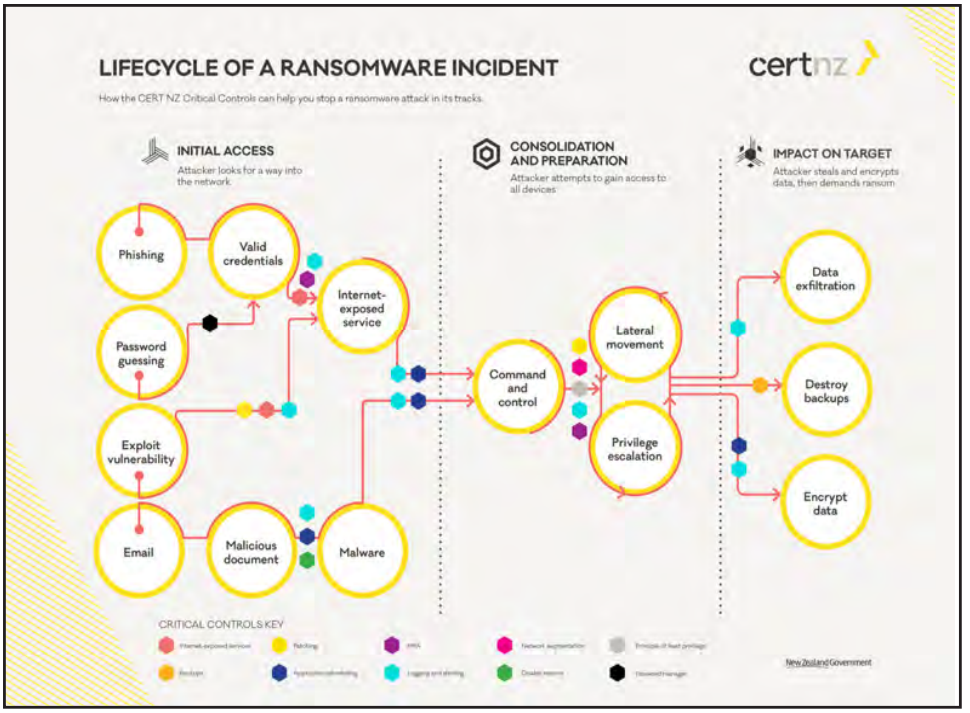


Figure 3-6. Life cycle of a ransomware attack
(Diagram by CERT NZ)

Even though figure 3-6 presents a ransomware campaign in three sections—initial access, consolidation and preparation, and impact on target—this model does not account for the negotiation process. To understand a ransomware incident from start to finish, it is helpful to examine the six major phases that comprise the life cycle of a ransomware attack, which has many similarities to the Cyber Kill Chain® discussed earlier in the chapter.

The first step of a ransomware attack is the distribution campaign, in which attackers seek potential victims to infiltrate relatively easily and with minimal time and effort. Cyber threat actors use several methods to find appropriate targets, such as vulnerability indexing search engines.[37] For instance, attackers using Shodan—a search engine for Internet-connected devices—or similar Internet scanning programs could find potential targets. As part of step one, the attacker tries to infect malicious files (payload) to the target in several ways, such as a phishing attack, a watering-hole attack, an exploit kit, or a drive-by-download.

37. Benjamin David, "More than Two-Thirds of Organizations Are Targets of at Least One Ransomware Attack," *InfoSecurity Magazine* (website), September 29, 2021, https://www.infosecurity-magazine.com/news/two-thirds-organizations-ransomware/.

Stage two, infiltration and staging, begins when the attacker discovers ways to access the targeted network. Here, the attacker tries to understand the network and the limitations of its account. Later, the attacker uses stealth and camouflages executable programs to understand the scene better. The staging phase is mainly related to checking the local configurations and seeking registry keys for various rights and proxy settings, user privileges, and accessibility. Attackers cautiously use Internet Protocol analytic tools to understand the capability of the target.

After infiltration and staging, the third stage is scanning, in which attackers focus on the details that will ensure the success of the attack. All IT and OT systems have their own peculiarities, but the business models and sectoral software in which they operate can change the settings. Attackers check the backup structure and critical files to halt the business activity, and inspect the target's financial condition to determine the ransom amount. Finally, attackers try to obtain administrative privileges to control the security systems, like the security information and event management tools, endpoint detection and response systems, and virtualization platforms.

In the fourth stage, encryption, attackers encrypt the target's files in priority order without being noticed by the antivirus software. Typically, attackers prioritize the files used in daily operations during the encryption process because these files tend to be essential for the continuity of the business. Since encryption is a time-sensitive race that the attackers try to complete as soon as possible—to prevent the target from using backups to restore the system—they often prefer weekends to start their encryption operations. The attacker also downloads critical data of the target system to be used as leverage in the next step of the attack cycle.

Step five, discovery and ransom demand, begins when the targeted business systems are disrupted or halted—often on the first day of the work week—and the business owner receives a note from the ransomware attack group. In most cases, this note is a text file on the desktop that includes a countdown timer, a list of frequently asked questions regarding the ransomware operation, and the attackers' demand for a specific ransom amount to decrypt the target's files.

The final phase, negotiation and settlement, starts when the target understands that its systems are paralyzed by a ransomware campaign. Identifying how the attackers breached the targeted system and learning which ransomware variant encrypted the data are vital elements to managing

the negotiations. Since the critical infrastructure sectors are also part of national security, notifying the law enforcement, relevant state institutions, and the insurance carrier are important initial steps. Communications with the attackers are essential to resolving the crisis, but before beginning direct negotiations the target should analyze all possible outcomes and alternative plans.

Once the targeted institution decides to begin negotiations, the organization should designate a lead negotiator and determine how to make decisions throughout the process. During the negotiations, the designated negotiator should demand some form of proof from the attack group—such as the system's directory plans and proof of decryption capability—to understand the ransomware group's intentions and capacity. The attackers, in most cases, search the target network in-depth, so they know the target's financial capacity to pay a ransom. The negotiator should ask for a discount before the final settlement between the targeted institution and the attack group, which is the last step of the process. After the settlement, the target should conduct a deep forensic analysis of the incident and take the necessary steps to prevent future cyberattacks.

According to a cyber alert issued by the US Cyber Security and Infrastructure Agency, an unnamed natural-gas compression facility based in the United States was the target of a ransomware attack in 2020. Using a spear-phishing link to obtain initial access to the organization's IT and OT networks, the attackers deployed commodity ransomware to encrypt data on these networks for maximum destruction and then requested a ransom payment. On the OT network, human machine interfaces, data historians, and polling servers all experienced loss of availability, a loss of real-time operational data that ultimately resulted in a partial loss of view for human operators.[38] Although the facility maintained control of its programmable logic controllers and overall operations, the facility's emergency response framework prioritized physical safety threats over cyber incidents. Thus, the facility management decided to stop its operations for two days.

38. "Alert (AA20-049A): Ransomware Impacting Pipeline Operations," Cyber and Infrastructure Security Agency (website), October 24, 2020, https://us-cert.cisa.gov/ncas/alerts/aa20-049a.

In summary, the ransomware attacks against this compression facility and the Colonial Pipeline discussed earlier in this section illustrate the recent increase in ransomware campaigns and the rising risks they pose to critical infrastructures.

Business E-mail Compromise (BEC)

Accompanying the rise in ransomware campaigns over the past few years is the increase in attacks on business e-mails, also known as e-mail account compromise. The FBI recorded a distinct increase in these types of attacks since 2015—with a sharp uptick in the monetary losses due to these attacks since 2017—to the cost of more than $26 billion through 2019 (see figure 3-7).[39]

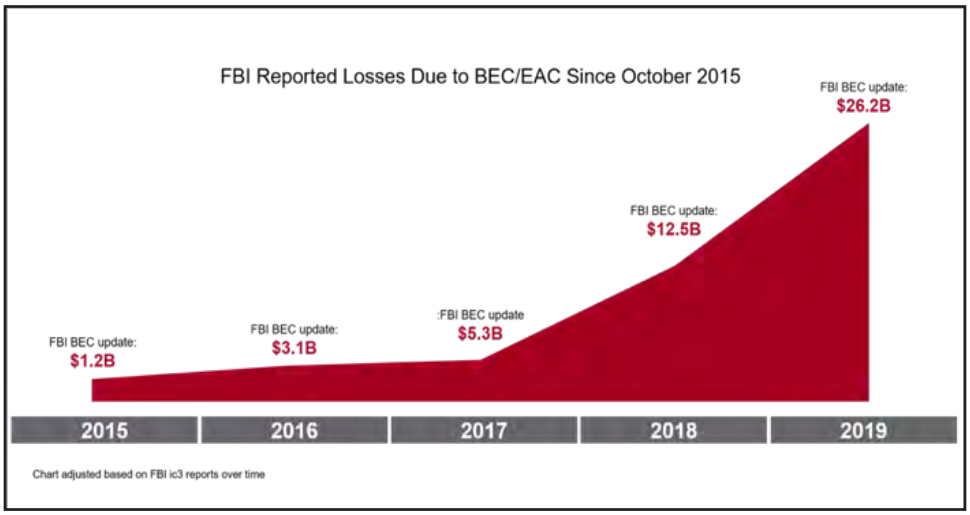


Figure 3-7. Annual losses due to business e-mail and e-mail account compromise
(Diagram by the FBI Internet Crime Complaint Center)

In 2020, the FBI Internet Crime Complaint Center received 19,369 complaints regarding business e-mail compromise.[40] The concept of BEC is a sophisticated scam that targets unsuspecting executives and employees into making payments or sending sensitive data to fraudulent accounts. The scam is frequently carried out by using techniques like social engineering or computer intrusion, which manipulate users into sending

39. Ryan Terry, "Business Email Compromise Results in $26B in Losses over the Last Three Years," *Proofpoint* (blog), September 12, 2019, https://www.proofpoint.com/us/blog/threat-protection/business-email-compromise-results-26b-losses-over-last-three-years.

40. "Internet Crime Report 2020," Federal Bureau of Investigation (website), n.d., accessed on September 5, 2021, https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf.

money or data. Since these attacks are highly targeted and include no payload, they are notoriously difficult to prevent. The established threat detection solutions that analyze e-mail headers, links, and metadata often miss these types of attacks. In a typical BEC example, HR or financial representatives receive e-mails that appear to be from higher-level employees, requesting them to update their direct-deposit information for the pay period. The financial information provided to HR or payroll representatives generally leads directly to the criminal's account. These types of attacks typically target the high-level management in a critical infrastructure environment and compel them to send technical or other information that could disrupt the critical infrastructure's services or lay the groundwork for a more destructive cyberattack in the future.

Credential Stuffing

Increasing digitalization brings with it the problem of authentication on various platforms. As a result, credential stuffing—a type of cyberattack that obtains compromised usernames and passwords to access user accounts—is increasing in frequency. Citizens of NATO member states and partner nations need numerous usernames and passwords to access state services, bank accounts or automated teller machines, e-mail and social media accounts, and a host of other functions for daily life. Given the limitations of human memory, people tend to reuse similar and predictable passwords for their accounts and services. The increase of data leaks and breaches on various sites has compromised a massive number of usernames and passwords. In a recent study on the use of passwords, results show that even after being notified that their personal data had been compromised, only about one third of users created new passwords, most of which were not strong or unique.[41] Another recent study indicated that five billion unique user credentials are circulating on dark net forums, where cybercriminals are selling bank accounts and domain administrator access credentials to other criminals in the different outlets of the dark web (see figure 3-8 for the 11 different categories of account listings by percentage).[42]

41. Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia, "(How) Do People Change Their Passwords after a Breach?," IEEE (website), n.d., accessed on September 12, 2021, https://www.ieee-security.org/TC/SPW2020/ConPro/papers/bhagavatula-conpro20.pdf.

42. Digital Shadows Photon Research Team, *From Exposure to Takeover: The 15 Billion Stolen Credentials Allowing Account Takeovers* (San Francisco: Digital Shadows, 2021), 2, 8, https://resources.digitalshadows.com/whitepapers-and-reports/from-exposure-to-takeover.

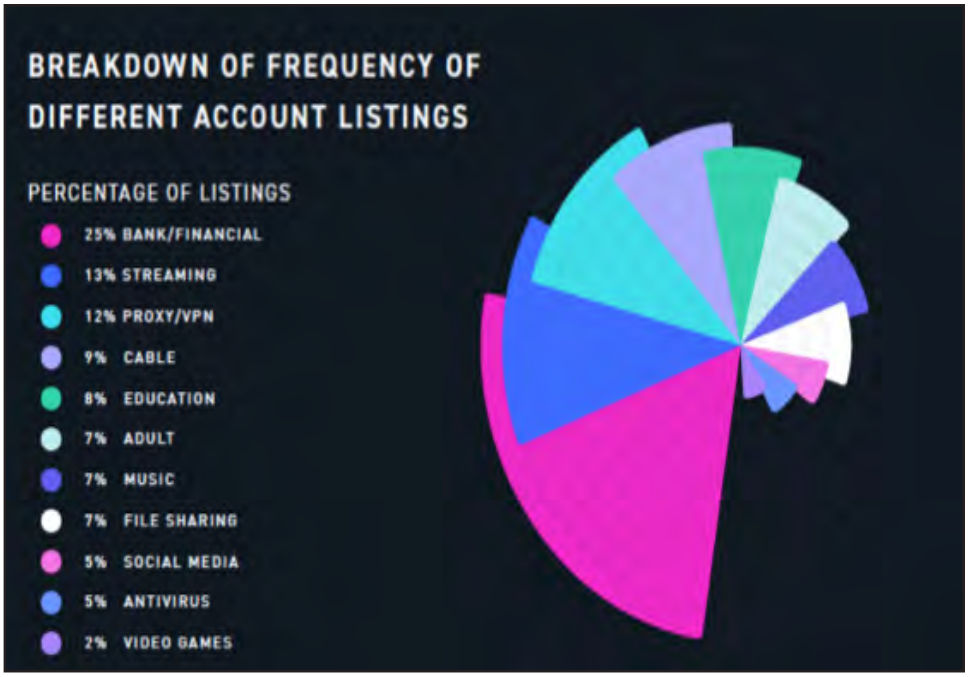


Figure 3-8. Listing of compromised accounts on the dark net
(Diagram by Digital Shadows)

In credential stuffing, an attacker sets up a botnet to log the compromised credentials into the platforms while simultaneously changing the Internet Protocol addresses. Attackers record the successful logins and either obtain identifiable information or money from this account or store it for future use. In the advanced form of the credential stuffing, the attackers conduct data scraping and scanning to sites like LinkedIn, Facebook, Twitter, and Amazon by using Open Bullet types of software to find real people and determine their affiliations.

A recent example of credential stuffing is the Colonial Pipeline ransomware attack. The attackers accessed Colonial Pipeline via a virtual private network (VPN) service, which was built to access the company's network remotely. At the time of the attacks, the VPN account was still functioning but not active. There is no specific information regarding where the attackers obtained these credentials. The VPN account's password, however, has been discovered among the leaked passwords on the dark web.[43] Another example is the cyberattack against the video surveillance startup Verkada, in which 150,000 cameras were compromised by a hacktivist collective known

43. William Turton and Kartikay Mehrota, "Hackers Breached Colonial Pipeline Using Compromised Password," Bloomberg (website), June 4, 2021, https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password.

as APT-69420 Arson Cats. The hackers initiated a credential-stuffing attack by gaining "super admin" level access to Verkada's system using a username and password they found in the public domain on the Internet.[44]

The results of the studies and examples of credential stuffing attacks cited in this section point to the reality that if a person has one compromised account, it can easily lead to multiple accounts being compromised. This point also blurs the distinction between personal and work accounts, and reveals the risks that employee password management practices, or lack thereof, can pose to critical infrastructure systems in the form of credential stuffing.

Supply Chain Attacks

Beyond the vulnerabilities and risks inherent to the human workforce, most critical infrastructures rely on supply chains with ever-changing cyber defense postures and moderate to high-level dependencies on external organizations. Supply chain attacks are a means to target and exploit legitimate trusted relationships between critical infrastructures and the external organizations that enable their operations. The ultimate goal of these attacks is to gain access to and compromise a vendor's systems, and then expand this access deeper into the affiliated organizations. Supply chain attacks typically include two significant stages. First, cyber-threat actors stealthily infiltrate the supplier company and its networks. The second step, to introduce the malware across the network, is much easier than the first one. Since the supply chain nodes and the supplier already have an established and trusting relationship, users tend to accept any updates and patches coming from the supplier.

A recent example of a supply chain attack is the SolarWinds attack in 2020, which shocked cybersecurity experts because it affected thousands of clients across the globe. SolarWinds's primary product, Orion, is a performance-monitoring platform to optimize the IT infrastructures of the companies who use the service, totaling some 300,000 customers. In the SolarWinds attack, cyberattackers used a known product (Orion) within a trusted relationship to compromise these companies and institutions, effortlessly gaining access to their protected systems. On December 13, 2020, a well-known cybersecurity firm, FireEye, released a report on the SolarWinds attack, noting the threat actors were conducting a global intrusion campaign using malware named SUNBURST. According to the report:

44. "150,000 Verkada Security Cameras Hacked—to Make a Point," *Malwarebytes Labs* (blog), March 12, 2021, https://blog.malwarebytes.com/iot/2021/03/150000-verkada-security-cameras-hacked-to-make-a-point/.

> FireEye has uncovered a widespread campaign that we are tracking as UNC2452. The actors behind this campaign gained access to numerous public and private organizations around the world. They gained access to victims via trojanized updates to SolarWind's Orion IT monitoring and management software. This campaign may have begun as early as Spring 2020 and is currently ongoing.[45]

It was later confirmed the attackers gained access to SolarWinds's Orion software in March 2020 and inserted a malicious code into the dynamic link library of Orion's update program. As soon as the companies using Orion updated the program—the first stage of this supply chain attack—this code activated a backdoor, which controls the setting of the network and transfers the necessary information to the command-and-control server of the cyber-threat actors. In the second stage, the threat actors gained an astonishing level of access, with capabilities such as privilege escalation and lateral movement (see figure 3-9 for an overview of the stages of the SolarWinds attack).[46]

Estimates indicate roughly 18,000 customers installed the Orion security updates in March 2020, meaning that some 6 percent of SolarWinds's customer base had infected systems and were vulnerable for the majority of 2020 prior to FireEye identifying the cyberattack.[47] Since the customer base consists of users from across the public and private sectors, the attack affected a number of high-profile organizations—such as AT&T, CISCO, McAfee, Microsoft, the *New York Times*, Symantec, and Visa—numerous universities, and US government agencies, including: the Departments of Commerce, Defense, Energy, Homeland Security, State, and Treasury; the Centers for Disease Control and Prevention; and the National Nuclear Security Administration.[48] This list is not exhaustive, and it is highly likely many more organizations suffered from the attack.

45. FireEye, "Highly Evasive Attacker Leverages SolarWinds Supply Chain to Compromise Multiple Global Victims with SUNBURST Backdoor," *Mandiant* (blog), accessed on July 23, 2021, https://www.mandiant.com/resources/blog/evasive-attacker-leverages-solarwinds-supply-chain-compromises-with-sunburst-backdoor.

46. Edward Cost, "What Is a Supply Chain Attack? Why You Should Be Worried about Your Vendors," *UpGuard* (blog), accessed on September 21, 2021, https://www.upguard.com/blog/supply-chain-attack.

47. Sam Ingalls, "FireEye, SolarWinds Breaches: Implications and Protections," eSecurity Planet (website), December 18, 2020, https://www.esecurityplanet.com/threats/fireeye-solarwinds-breaches-implications-protections/.

48. Ingalls, "FireEye, SolarWinds Breaches."

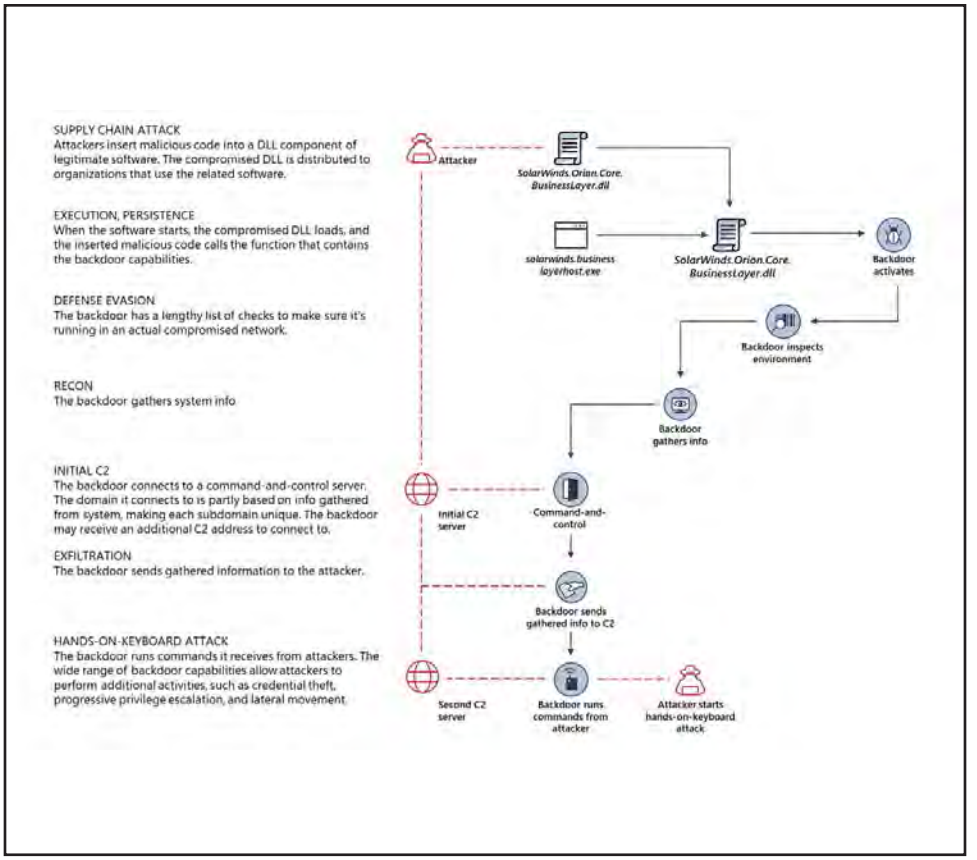


Figure 3-9. SolarWinds supply chain attack operation
(Diagram by Microsoft)

Just five days before it warned the public about the SolarWinds attack, FireEye announced that "a nation with top-tier offensive capabilities" had infiltrated its network and stolen the company's suite of red team tools, which FireEye used to mimic potential adversary attack and exploitation capabilities as it developed more effective cybersecurity products.[49] In response to the SolarWinds attack, which gave the threat actors access to FireEye's high-end security tools as well as sensitive information and critical cybersecurity infrastructure, analysts initially pointed to Russia as the source of the cyberattack. It took several months to collect the forensic evidence from the SolarWinds attack and understand the damage it had caused. Given the challenges in the attribution of cyberattacks, US authorities were initially cautious about announcing what group carried

49. FireEye, "FireEye Red Team Tools," Mandiant (website), December 8, 2020, https://www.mandiant.com/resources/unauthorized-access-of-fireeye-red-team-tools; and Ingalls, "FireEye, SolarWinds Breaches."

out the attack. In April 2021, however, the National Security Agency, the Cybersecurity and Infrastructure Security Agency, and the FBI released a joint cybersecurity advisory attributing the SolarWinds attack—and several others as well—to Russian Foreign Intelligence Service actors known as APT29, Cozy Bear, and The Dukes.[50]

Conclusion

The role of cybersecurity within CISR efforts presents particular distinctions and difficulties in several key areas. Dynamics such as rapidly changing technology, the nature and challenges of private-public cooperation, business-minded security investments, varied levels of experience among the workforce, and supply-chain problems combine to create a hyper-competitive landscape. These challenges simply illuminate the stark reality that there are limits to what and how much stakeholders can truly secure in critical infrastructure. Vulnerabilities change, threats evolve, and maintaining the necessary human capital in the workforce becomes increasingly difficult from year to year. The conflicts and frictions in the physical domain also exist in the cyber domain.

To sustain operations, government agencies try to regulate all possible aspects of critical infrastructure by focusing on the threats posed by malevolent forces. Well-written regulations, however, do not automatically address the requirements and expectations of the various critical infrastructure sectors nor do they facilitate much-needed technological improvements. The element of uncertainty also hinders the psychology required to make changes and preparations to enhance resilience. Additionally, cybersecurity teams are often severely short of personnel, and they suffer from limited training time to complete regular certifications and thus adapt to new technological developments and practices used by adversaries. Under these conditions, it is laborious to pursue CISR efforts with a traditional security mindset. Limited resources require new perspectives and practices.

50. "Cybersecurity Advisory: Russian SVR Targets U.S. and Allied Networks," Cybersecurity and Infrastructure SecurityAgency, April 2021, https://media.defense.gov/2021/Apr/15/2002621240/-1/-1/0/CSA_SVR_TARGETS_US_ALLIES_UOO13234021.PDF/CSA_SVR_TARGETS_US_ALLIES_UOO13234021.PDF.

The question of what mindset to adopt and how to do it, however, is one of the most challenging questions facing critical infrastructure stakeholders. In a recent interview, cybersecurity expert and business executive Dmitri Alperovitch offers a potential way ahead in forming a new cybersecurity perspective:

> I've been saying for 10+ years that intrusions are inevitable, no one is immune, and everyone needs to start thinking about this in terms of "we will likely get breached, we will likely get compromised, but how do we stop the damage from being done?" . . . The right way to think about security strategies going forward is to assume a breach, hunt continuously for any presence of adversaries on your network, and kick them out as quickly as possible.[51]

The cyber domain—and the increasing connectedness it brings to nearly all aspects of modern life—offers few clear borders or lines of defense to keep adversaries away. This situation creates an ambiguity that requires critical infrastructure stakeholders to adopt new perspectives and implement new CISR policies that can survive and succeed under these conditions. The best possible alternative as the way forward to building more resilient and secure systems is to adopt a zero-trust mentality and "assume breach" mindset for the future.[52] This new security mindset should focus on investment in human capital and improving the training, situational awareness, and overall capability of the critical infrastructure cybersecurity workforce.

51. Dmitri Alperovitch, "SolarWinds Breach: An RSAC Interview with Dmitri Alperovitch about Who, How and Why," RSA Conference, December 14, 2020, YouTube video, 6:46, https://www.youtube.com/watch?v=3kpaV4FNzc0.

52. Microsoft Cyber Defense Operations Center, "Strategy Brief," Microsoft Corporation (website), n.d., accessed on September 9, 2021, https://download.microsoft.com/download/4/6/8/4680DFC2-7D56-460F-AD41-612F1A131A26/Microsoft_Cyber_Defense_Operations_Center_strategy_brief_EN_US.pdf.

— 4 —

Hybrid Threats to US and NATO Critical Infrastructure

Carol V. Evans

Protecting critical infrastructure is a vital, strategic security concern and challenge for the United States and the North Atlantic Treaty Organization. Adversaries are actively targeting critical infrastructure in the Allied member states—particularly the energy, transportation, information, communications, and the defense industrial base (DIB) sectors—as a potential means to undermine military capability, force projection, mobility, and sustainment.

To provide an understanding of adversarial hybrid threats to critical infrastructure and the innovative ways in which the United States and NATO are countering them, this chapter is structured in three sections. The first section addresses the evolution in the nature of the threat to critical infrastructure. The technological convergence between communications and information technology, with cyber connectivity to critical infrastructure systems, has shifted the threat paradigm from kinetic to a cyber and hybrid means of attack. This relatively recent development has created opportunities for adversaries to exploit vulnerabilities in the critical infrastructure upon which US and NATO armed forces depend. Section two provides an analysis of several hybrid threat vectors with the potential to attack, undermine, or compromise US and NATO warfighting, force projection and sustainment capabilities. The first vector contains hybrid threats to the US homeland—in particular, the deliberate infiltration of the energy infrastructure which supports US installations and bases to interfere

with the military's ability to deploy and sustain forward combat forces and equipment. A second hybrid threat vector is adversarial targeting of US and NATO logistics, with the potential to degrade US overseas force projection as well as NATO mobility and sustainment within the theater. The third hybrid threat stems from China's strategic penetration, ownership, and control of key DIB infrastructure and supply chains in Europe via its Belt and Road Initiative and foreign direct investment activities. This vector provides an opportunity to undermine US and NATO interoperability and political unity. Finally, the chapter concludes by highlighting US and NATO measures to redress and mitigate these threats by investing in critical infrastructure security and resilience through organizational capacity building, policy frameworks, and implementation of host country baseline resilience requirements.

Kinetic-Cyber-Hybrid Threats to Critical Infrastructure

The nature of the threat to the critical infrastructure in the United States and NATO countries has evolved significantly from one that was based primarily on kinetic attacks by terrorist organizations to the exploitation of cyber and hybrid means by nation-states, proxies, and other adversaries. Beginning in 2001, there was a rapid escalation of high-profile attacks involving critical infrastructure in Allied countries by al-Qaeda and other terrorist organizations. These include the 9/11 attacks against the Pentagon and World Trade Center, the 2004 Madrid commuter and Atocha train station bombings, the 2005 London transport bombings, the coordinated series of terrorist attacks in Paris in 2015, and the Atatürk airport shootings and suicide bombings in 2016. All of these examples required the attackers to be present physically and to use kinetic means (see chapter 2 for an overview of physical terrorist threats and chapters 6 and 7 for greater detail on these examples and others against the civil aviation and rail sectors).

The emergence of cyberspace, however, has provided a new delivery mechanism that has increased the speed, frequency, and power of attack against US and NATO critical infrastructure. The effectiveness of cyber tools has been augmented by the rapid interconnectivity of information and communication systems with critical infrastructure systems, including the rollout of the Internet of Things and the creation of smart cities (see chapters 3 and 14 for in-depth discussion on cyber threats and actors, and recommendations for managing cybersecurity risks). As the then US Director of National Intelligence, Admiral Dennis Blair, observed, "The growing connectivity between information systems, the Internet, and other infrastructures

creates opportunities for attackers to disrupt telecommunications, electrical power, energy pipelines, financial networks, and other critical infrastructure."[1] This information-communications-digital-cyber revolution—consisting of the interdependent networks of informational technology infrastructures (hardware, software, data, and protocols) and information (the Internet, telecommunications networks, computer systems, and embedded processors and controllers)—created an attack vector by which the cyber realm has now become the favored, go-to weapon by adversaries of the United States and NATO. For adversaries, the cyber domain provides anonymity and lower risk of detection and personal injury, and requires few resources to access a wide range of diverse targets—all with the ability to operate from nearly any geographic location. Hence, the proliferation in cyberattacks against US and NATO infrastructure, as it is possible to attack strategic targets with minimal exposure, without physically being present or having to confront defensive forces.

These attacks are increasing in regularity and sophistication by China, Russia, Iran, and North Korea. According to US intelligence sources, China "presents a persistent cyber . . . attack threat to our core military and critical infrastructure systems," while Russia "poses a cyber espionage, influence, and attack threat to the United States and our allies," and ". . . is now staging cyber-attack assets to allow it to disrupt or damage US civilian and military infrastructure during a crisis."[2] Operation Cleaver—an extensive, global surveillance and infiltration campaign—has been attributed to Iran. The targets of this campaign included oil and gas, energy and utilities, mass transportation, airlines, airports, hospitals, telecommunications, DIB, chemical companies, and government and military networks in NATO countries: Canada, France, Germany, Turkey, the United Kingdom, and the United States.[3] Among the world's most sophisticated threat groups, North Korea's Reconnaissance General Bureau—commonly known as the Lazarus Group or APT38—has launched spear-phishing attacks against employees of American energy, aerospace, and technology companies, as well as the US Departments of State and

1. Director of National Intelligence Admiral Dennis C. Blair, *Annual Threat Assessment of the Intelligence Community for the Senate Armed Services Committee: Statement for the Record*, March 10, 2009, 39–40.

2. Daniel R. Coats, *Worldwide Threat Assessment of the US Intelligence Community: Statement for the Record, Senate Select Committee on Intelligence* (Washington, DC: Office of the Director of National Intelligence, January 29, 2019), 5, https://www.dni.gov/files/ODNI/documents/2019-ATA-SFR---SSCI.pdf.

3. Tony Bradley, "Cylance Unveils Details of Iran-based Hacking in 'Operation Cleaver' Report," CSO (website), December 3, 2014, https://www.csoonline.com/article/2854686/cylance-unveils-details-of-iran-based-hacking-in-operation-cleaver-report.html.

Defense. The group was also responsible for the 2017 WannaCry ransomware attack, which brought the UK National Health Service to a halt.[4] In 2021, the US Department of Justice charged three North Korean computer programmers affiliated with APT38 with a range of high-profile cyberattacks. Of note, North Korean cyber teams have recently targeted financial institutions and virtual currency exchanges to generate funds to support the country's ballistic missile program.[5]

Russia's invasion and annexation of Crimea in 2014 and support for a separatist insurgency in eastern Ukraine led to the emergence of "hybrid warfare" and other terms such as "grey zone conflict," and "unrestricted warfare" to describe what was considered to be a new form of warfare.[6] Hybrid warfare also has been applied to Russian activities in Georgia and in the Baltic countries of Estonia and Latvia (both members of NATO and the European Union). Such activities triggered concern "that Russia will seek to use the Russian minority to gain influence in the Baltics, use covert action to seize territory, use subversion to justify a conventional attack, or otherwise use deniable or covert means to gain influence in the Baltics and undermine the EU and NATO."[7] Implicit in these hybrid warfare examples is the inherent ambiguity of Russian actions that would impede a timely and coordinated response from NATO, thereby undermining its credibility and commitment to the eastern Allies, and effectively giving Russia veto power over Euro-Atlantic enlargement.

The concept of hybrid warfare has spawned a lively debate in recent years among military strategists and academics, with many experts contending that there is nothing new in this type of warfare and that its analytic utility

4. Graham Cluley, "US Charges North Koreans in Relation to Global Cyber Attacks," Tripwire (website), February 18, 2021, https://www.tripwire.com/state-of-security/featured/us-charges-north-korean-hackers-wannacry-sony-pictures-attack/.

5. Edith M. Lederer, "UN Experts: North Korea Using Cyber Attacks to Upgrade Nukes," AP News (website), February 9, 2021, https://apnews.com/article/technology-global-trade-nuclear-weapons-north-korea-coronavirus-pandemic-19f536cac4a84780f54a3279ef707b33.

6. Matthew Kofman and Michael Rojansky, *A Closer look at Russia's "Hybrid War"* (Washington, DC: Wilson Center, April 2015), 1–2, https://www.wilsoncenter.org/sites/default/files/media/documents/publication/7-KENNAN%20CABLE-ROJANSKY%20KOFMAN.pdf; and Keir Giles, *Russia's 'New' Tools for Confronting the West: Continuity and Innovation in Moscow's Exercise of Power* (London: Royal Institute of International Affairs, March 2016), 7–9, https://www.chathamhouse.org/2016/03/russias-new-tools-confronting-west-continuity-and-innovation-moscows-exercise-power.

7. Andrew Radin, *Hybrid Warfare in the Baltics: Threats and Potential Response* (Santa Monica, CA: Rand Corporation, 2017), 1.

should be contested.[8] Indeed, NATO has moved away from the use of hybrid warfare to the larger construct of hybrid threats. Under this rubric, hybrid warfare is a component in a range of hybrid threat activity. The latter is not expected to trigger Article 5 of the Washington Treaty. According to NATO, hybrid threats:

> combine military and non-military as well as covert and overt means, including disinformation, cyberattacks, economic pressure, deployment of irregular armed groups and use of regular forces. Hybrid methods are used to blur the lines between war and peace, and attempt to sow doubt in the minds of target populations. They aim to destabilize and undermine societies.[9]

In this respect, the hybrid threat concept underscores the systemic vulnerabilities of democratic states by revisionist adversaries and authoritarian states, which require countermeasures involving a whole of government approach and civil-military cooperation.[10]

Hybrid threats provide a useful framework to understand why critical infrastructure is increasingly weaponized by US and NATO adversaries. Infrastructures are attractive targets to coerce, intimidate, and apply pressure on a target state, as demonstrated by Russian cyberattacks against Ukraine's electric grid in 2015 and 2016, and against a range of critical infrastructure in the days and hours before Russia's invasion of Ukraine in February 2022. Adversaries can also use kinetic means to achieve those effects without necessarily engaging in open military activity. A good example of this type of hybrid threat occurred in 2017, when Russian naval ships hindered the installation of an undersea electricity cable between Sweden and Lithuania by obstructing the Nordbalt cable-laying vessels. The power cable was laid to increase energy supply in both countries and to facilitate the exchange of power between the Baltic and Nordic electricity markets, thereby decreasing their combined

8. Antulio J. Echevarria II, *Operating in the Gray Zone: An Alternative Paradigm for US Military Strategy* (Carlisle, PA: Strategic Studies Institute, US Army War College Press, 2016), 1, https://press.armywarcollege.edu/monographs/425/.

9. "NATO's Response to Hybrid Threats," NATO (website), March 16, 2021, https://www.nato.int/cps/en/natohq/topics_156338.htm#:~:text=Hybrid%20threats%20combine%20military%20and,and%20use%20of%20regular%20forces.

10. European Union and Hybrid Centre of Excellence, *The Landscape of Hybrid Threats: A Conceptual Model* (Luxembourg: Publications Office of the European Union, 2021), 9, https://www.hybridcoe.fi/publications/the-landscape-of-hybrid-threats-a-conceptual-model/.

energy dependence on Russia.[11] The next section will showcase how adversaries use myriad hybrid threat vectors to create and then leverage civil-military infrastructure dependencies in ways that have deleterious impacts to US and NATO military power.

Prepping the Battlespace: Weaponizing Critical Infrastructure to Challenge US and NATO Military Supremacy

The deterrent value of US and NATO forces is critical. Deterrence is based not just on credible military capabilities, force structure, and force projection but on the ability of critical infrastructure in the United States and NATO member states to support short-fused response, reinforcement timelines, and means of sustainment. Penetrating, disrupting, controlling, and destroying key global critical infrastructure is an advanced instrument by adversaries to degrade NATO missions and operations. This section will examine three areas where adversaries are deploying hybrid threats against critical infrastructure in the United States and Europe to undermine such supremacy.

Hybrid Threats to the US Homeland and Warfighting Capabilities

Adversaries—in particular Russia and China—are deliberately and effectively targeting the energy infrastructure, especially US electric grids, necessary to support US military installations and bases to compromise future warfighting capabilities. The electric grid is the key lifeline sector that powers all other civil-military infrastructure sectors: water, sanitation, communications, and transportation. Kinetic attacks on the US power grid have been considered by Russia and China within the context of preemptive first-strike capabilities.[12] Cyberattacks, however, are the primary means to strike US—and NATO host country—grids, because modern power grid interconnections rely on complex supervisory control and data acquisition (SCADA) systems, as well as use of new communication and network technologies, that provide back door access for potential adversaries. It is instructive to briefly examine Russia's cyberattacks on Ukraine's electricity grid as these attacks enabled Moscow to test, prove, and refine

11. Alexandra Brzozowski, "NATO Seeks Ways of Protecting Undersea Cables from Russian Attacks," Euractiv (website), October 23, 2020, https://www.euractiv.com/section/defence-and-security/news/nato-seeks-ways-of-protecting-undersea-cables-from-russian-attacks/.

12. *Terrorism and the EMP Threat to Homeland Security: Hearing before the Committee on the Judiciary of the United States Senate, Subcommittee on Terrorism, Technology, and Homeland Security*, 109th Cong. (2005) (statement of Dr. Peter Pry, Congressional EMP Commission Senior Staff), https://www.govinfo.gov/content/pkg/CHRG-109shrg21324/html/CHRG-109shrg21324.htm.

its cyber warfare capabilities for future employment in the United States (see chapter 5 for more thorough examination of these cyberattacks).

The linkage between critical infrastructures as an instrument of hybrid warfare has been on open display in Ukraine, where a Russian cyber army closely affiliated with the Kremlin, has systematically attacked almost every sector of Ukraine's infrastructure since 2015.[13] These attacks were set against the backdrop of Russia's illegal annexation of Crimea in 2014 and continued military clashes in the eastern Donetsk and Luhansk regions in Ukraine. Most notable were the attacks against Ukraine's electric grid in December 2015—which left large parts of the capital city, Kiev, and the western region of Ivano-Frankivsk in the dark—followed by another, more technologically sophisticated, attack in 2016 against one of Kiev's transmission substations.[14] These cyberattacks were attributed to a Russian group known as Sandstorm, which deployed its BlackEnergy malware to penetrate specialized computer architectures that are used for remotely managing physical industrial equipment and control systems. What was most worrying to cyber experts was that Sandstorm had already targeted NATO networks, and had compromised the computers of American and European electric and water utility companies with the same Trojan malware. This malware provided the group with enough control to induce blackouts on US soil. As one cyber forensic expert forewarned, "An adversary that had already targeted American energy utilities had crossed the line and taken down a power grid [in Ukraine]. It was an imminent threat to the United States."[15]

In March 2018, the US Federal Bureau of Investigation and the Department of Homeland Security confirmed that Russian government cyber hacker teams had actively "targeted government entities and multiple US critical infrastructure sectors, including the energy, nuclear, commercial facilities, water, aviation, and critical manufacturing sectors."[16] The Russian cyberattack teams included Sandstorm, Dragonfly, and Palmetto Fusion, with some attributed to gaining remote access to actual industrial control systems and US energy sector networks, including a Kansas nuclear power

13. Andy Greenberg, "How an Entire Nation Became Russia's Test Lab for Cyberwar," Wired (website), June 20, 2018, https://wired.com/story/russian-hackers-attack-ukraine/.

14. Kim Zetter, "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid," Wired (website), March 3, 2016, https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/.

15. Andy Greenberg, *Sandstorm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (New York: Doubleday, 2019), 53.

16. "Alert TA18-074A: Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure," Cyber and Infrastructure Security Agency (website), March 15, 2018, https://www.us-cert.gov/ncas/alerts/TA18-074A.

facility.[17] Cyberattacks against the US power grid have continued. For instance, the threat group known both as Xenotime or Triton has compromised electric facility safety systems to cause potential plant disruption and damage. According to a researcher at the US cybersecurity firm Dragos, surveillance of the US electric grid is "indicative of the preliminary actions required to set up for a future intrusion and potentially a future attack."[18]

Cyberattacks have not been limited to Russian perpetrators only. China is also assessed to have the ability to shut down the US power grid through a cyberattack. In 2014, the then Commander of US Cyber Command and Director of the National Security Agency, Admiral Michael Rogers, testified that the United States had detected malware from China on US systems, which enabled Beijing "to shut down very segmented, very tailored parts of our infrastructure."[19] Mandiant, a cybersecurity firm, also confirmed that hackers working on behalf of the Chinese government were actively penetrating American public utility systems that service everything from power generation to the movement of water and fuel across the country.[20]

Hybrid threats against US energy infrastructure have come at a time when a number of factors leave this infrastructure more vulnerable. In the United States, most electricity consumed by military installations is sourced from the commercial power grid. In recent years, these military bases have been used to conduct specialized warfighting activities, intelligence processing, exploitation and dissemination, and networked real-time communications for command and control. These activities have greatly increased US military dependence on energy, particularly electric power.

Decades-long underinvestment in US base facilities—combined with this increased reliance on privately owned infrastructure largely outside of the military's control—is a major compounding factor that provides another source for adversarial exploitation. Increasing geographic

17. John Kennedy, "US Officially Blames Russia's 'Dragonfly' Hackers for Attacks on Energy Grid," Silicon Republic (website), March 26, 2018, https://www.siliconrepublic.com/enterprise/dragonfly-us-russia-energy-grid-hackers.

18. Andy Greenberg, "The Highly Dangerous 'Triton' Hackers Have Probed the US Grid," Wired (website), June 14, 2019, https://www.wired.com/story/triton-hackers-scan-us-power-grid/.

19. *Cybersecurity Threats: The Way Forward: Hearing before the House Select Committee on Intelligence*, 113th Cong. (2014) (statement of Admiral Michael Rogers, Commander, US Cyber Command and Director, National Security Agency), https://www.nsa.gov/Press-Room/Speeches-Testimony/Article-View/Article/1620360/hearing-of-the-house-select-intelligence-committee-subject-cybersecurity-threat/.

20. Jamie Crawford and National Security Producer, "The U.S. Government Thinks China Could Take Down the Power Grid," *CNN* (website), November 21, 2014, https://www.cnn.com/2014/11/20/politics/nsa-china-power-grid/index.html.

concentration of infrastructure is one such factor. For example, over 31 percent of US naval ship-building and repair capacity is in and around Norfolk, Virginia.[21] Department of Defense (DoD) installations and associated infrastructure depend on continuous, assured power to support missions and operations both in the continental United States and abroad. Any extended loss of power is what has been acknowledged as a glaring national security Achilles' heel. America must expect adversaries to target and attempt to disrupt its power grid, with potential cascading and escalating failures in transportation, telecommunications, and other critical infrastructure services upon which the US military depends (see chapter 12 for further explanation of these interdependencies and potential failures as well as the necessity of building resilience strategies that sufficiently account for them).

As one former senior DoD official conceded, "The smart thing to do is to maneuver around those forces, attack the critical infrastructure, the facilities here in the United States on which we depend to deploy, operate and sustain our forces abroad."[22] The willingness and ability of adversaries to deploy destructive cyber weapons in future warfare with the United States has immense national security implications. Of immediate concern is the threat to US deterrence and intrinsic force projection capabilities, for: "[i]t does not matter how capable, how well trained or how advanced a nation's forces are if they can't get to the front in time."[23]

Hybrid Threats to US and NATO Mobility and Sustainment Operations

> Our deterrence and defense posture is underpinned by credible forces, both in-place and ready for reinforcement within Europe and from across the Atlantic.
>
> —2018 NATO Brussels Summit Declaration

21. Paul W. Parfomak, *Vulnerability of Concentrated Critical Infrastructure: Background and Policy Option*s, Congressional Research Service (CRS) Report PL33206 (Washington, DC: CRS, September 12, 2008), 4, https://crsreports.congress.gov/product/details?prodcode=RL33206.

22. Paul Stockton, quoted in Cynthia E. Ayers and Kenneth D. Chrosniak, *Terminal Blackout: Critical Electric Infrastructure Vulnerabilities and Civil-Military Resiliency*, Issue Paper 1-13 (Carlisle, PA: Strategic Studies Institute, US Army War College Press, October 2013), 5.

23. Omar Lamrani, "Why Logistics Will Be Key to Any U.S. Conflict with Russia and China," Rane Worldview (website), December 17, 2018, https://worldview.stratfor.com/article/why-logistics-will-be-key-any-us-conflict-russia-and-china/.

> NATO . . . must invest in its ability to monitor and defend against any Chinese activities that could impact collective defense, military readiness or resilience in the Supreme Allied Commander Europe's (SACEUR) Area of Responsibility.
>
> —NATO 2030: United for a New Era

The ability of the United States and NATO to project forces depends on mobility and sustainment, the latter of which is underpinned by secure and robust logistics. As the United States and NATO gear up to meet great power competition with Russia and China, strategic mobility to project and sustain military forces into these respective theaters will be critical. Further, strategic mobility is at the heart of a credible deterrence posture and, in this new hybrid threat environment, the capacity to project forces is vulnerable. Mobility begins with military facilities in the continental United States and does not end until the troops and equipment are postured forward and prepared for operations in theater. Indeed, the ability of the United States to mobilize its military capability for rapid response to crises, conflicts, and wars is what makes NATO credible as well.

This section analyzes how adversaries are identifying key US and NATO war-supporting logistics infrastructure and mobilization nodes to disrupt the timely preparation, deployment, and sustainment of military forces and material from the US homeland, as well as within NATO's member states. It first examines how adversaries may use hybrid threats to disrupt the projection of US military forces beyond North America and degrade the ability of the United States to conduct a sustained NATO war effort by targeting critical infrastructure.

Adversarial targeting of war-supporting infrastructure is not new. The Soviet Union understood that military personnel and supply reinforcements from the US homeland would prolong any future conflict in the European theater, resulting in a favorable disposition for NATO. Accordingly, Soviet military and intelligence establishments focused on damage to the US mobilization base and supporting infrastructures, including US military facilities and naval bases, commercial ports, railways, other transportation nodes, and lifeline infrastructure sectors. The Soviets conducted assessments of how US forces prepared for strategic deployment from the United States, what they mobilized, what kind of resources were required to transport and sustain deploying troops, and what military and civilian entities were involved

in that coordination.[24] An open source modeling assessment conducted by Argonne National Laboratory highlighted that successful attacks on US critical infrastructure would have serious military disruption impacts including longer movement times, effective loss of assets, and shortfalls in critical skills and material.[25]

While the then Soviet Union mapped out the US DIB logistics infrastructure for kinetic targeting purposes, today's adversaries have the capability to carry out attacks through hybrid threat modes. This ability means the US homeland can no longer be considered a sanctuary, and adversaries have the strategic reach and means to attack key nodes and dislocate force projection activities. Today's contested deployment environment is one in which national security policymakers and military planners can no longer assume logistics can arrive in the European theater unchallenged.

A contested environment becomes even more problematic given that forward-deployed US forces in Europe—with their rotational and force posture rebalancing requirements—are insufficient, which places an even larger premium on the need for rapid reinforcements of troops, munitions, and material. Yet, deploying military capabilities while under attack is a scenario for which the US defense and homeland security establishments remain largely underprepared.[26] Indeed, a recent study suggests that since the United States has not had "a comprehensive strategy to protect its civilian population and defense industrial base, or to mobilize and sustain the nation during time of war" since 1993, "America risks losing its next war with one or more major nation states."[27]

Compounding these challenges are significant critical weaknesses in US transportation and logistics infrastructure that undermine its force projection and sustainment capabilities and provide additional threat vectors to be exploited by adversaries. One such weakness is a near-term, crippling shortfall in strategic mobility which has been recognized as an unacceptable risk in force projection. At present, US Transportation

24. See Graham H. Turbiville Jr., "Prototypes for Targeting America: A Soviet Assessment," *Military Review* 82, no. 1 (January/February 2002): 3–9.

25. John Hummel, James F. Burke Jr., and William B. Cunningham, "Modeling the Impacts on National Security from Disruptions in CONUS Critical Infrastructures" (presentation, 72nd Military Operations Research Society Symposium, US Military Academy, West Point, NY, June 21, 2005).

26. See Bert Tussing and Barrett Parker, "The Multi-Domain Battle: What's in it for the Homeland?" War Room (website), November 10, 2017, https://warroom.armywarcollege.edu/articles/multi-domain-battle-whats-homeland/.

27. H. Quinton Lucie, "How FEMA Could Lose American's Next Great War," *Homeland Security Affairs* 15 (May 2019): 2, https://www.hsaj.org/articles/15017.

Command (USTRANSCOM)—the functional combatant command responsible for providing air, land, and sea transportation to meet US mobility needs—has sufficient transport aircraft to lift only one armored brigade combat team, with its roughly 5,000 troops and several hundred military vehicles, to a theater of operations. The picture is equally dismal for US strategic sealift capacity, as 70 percent of the organic fleet will be over 60 years old in 2034.[28]

In addition to an aging organic sealift fleet, TRANSCOM acknowledged that a reduction in US-flagged vessels, insufficient naval escorts, and a dwindling merchant marine fleet and seamen, are other competing, deleterious factors to strategic sealift capacity.[29] In 2018, the US Army warned the House Armed Services Committee that the nation's surge sealift capacity—which the Army and Marine Corps would rely on to transport up to 90 percent of their equipment in support of a major war or crisis—would fall below its requirement by 2024.[30] The Army noted in its information paper to the Committee that these sealift capacity shortfalls "undermine the effectiveness of US conventional deterrence as even a fully-resourced and trained force has limited deterrent value if an adversary believes they can achieve their strategic objective in the window of opportunity before American land forces arrive."[31]

TRANSCOM's mobility mission is heavily reliant on private sector commercial air, ground, and maritime transportation providers, which are very vulnerable to cyberattacks and energy disruptions. In his 2018 Senate testimony, General Darren McDew, the then commander of TRANSCOM, noted that its "logistics enterprise is more susceptible to these malicious [cyber] activities than other military organizations based on our unique relationships with commercial partners."[32] This vulnerability is especially apparent as 90 percent of TRANSCOM's military logistics and global operations are executed on unclassified commercial networks.[33] The Russian-launched

28. Lamrani, "Why Logistics Will Be Key."

29. *Posture of the United States Transportation Command: Hearing before the Senate Committee on Armed Services*, 115th Cong. (2018) (statement of General Darren W. McDew, Commander, US Transportation Command), 10, https://www.armed-services.senate.gov/hearings/18-04-10-posture-of-the-united-states-transportation-command.

30. David B. Larter, "US Army Warns of Crippling Sealift Shortfalls during Wartime," *Defense News* (website), November 11, 2018, https://www.defensenews.com/naval/2018/11/12/us-army-warns-of-crippling-sealift-shortfalls-during-wartime.

31. Larter, "US Army Warns."

32. *Posture of United States Transportation Command*, 18.

33. *Posture of United States Transportation Command*, 18.

NotPetya cyberattack in 2017 is instructive on potential future risks to TRANSCOM. This cyberattack, which originally targeted Ukraine, had the unintended effect of bringing one of TRANSCOM's major transport providers, Maersk, and its entire global port, shipping, logistics, and container operations to a halt for more than 10 days. Over 76 ports and 800 vessels were affected and cost Maersk an estimated $300 million to bring back its network systems.[34] The disruption to global supply chains caused by the NotPetya attack were enormous. Ssee chapter 5 for more detail on the nature and impacts of the NotPetya cyberattack.

Hybrid Threats from the People's Republic of China

Turning to the threats to NATO mobility and sustainment operations in the European theater, increasing analysis needs to be directed at Chinese hybrid activities and, in particular, to General Secretary of the Chinese Communist Party Xi Jinping's signature Belt and Road Initiative (BRI). Significant investments in Europe's maritime infrastructure have been channeled through the BRI's twenty-first-century Maritime Silk Road, which consists of three blue water passages and massive investments in port infrastructure, including deep sea ports and facilities, industrial free-trade zones, energy storage, pipelines, and refining facilities. The People's Republic of China (PRC) contends that the Maritime Silk Road is needed to secure its sea lanes of communication and trade routes for its energy, natural resources, and supply chain needs.

Two of the Maritime Silk Road's blue water passages, the China-Indian Ocean-Africa-Mediterranean passage and the China-Europe-Arctic Ocean route, are being developed to link and expedite trade and investment from China into Europe, the PRC's largest trading partner. Within Europe itself, China has rapidly expanded its port facility and terminal operations portfolios. Major Chinese port infrastructure projects include the Italian ports of Trieste, Venice, and Ravenna; the Slovenian port in Capodistria; and the Croatian port in Fiume. Additionally, Chinese state-owned companies—led by China Ocean Shipping Company (COSCO), the world's fourth largest container shipping fleet—have acquired controlling and minority stakes in 13 European ports. In Zeebrugge, Belgium's second largest port, COSCO owns 90 percent of the country's only terminal operator. In Spain, COSCO has a 51 percent stake in and managerial control of the largest terminal in Valencia as well as a 40 percent stake in Noatum

34. Andy Greenberg, "The Untold Story of NotPetya, The Most Devastating Cyberattack in History," Wired (website), August 22, 2018, https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world.

Container Terminal in Bilbao. It also has minority stakes in other terminals in Antwerp, Las Palmas, and Rotterdam. Together these ports account for roughly 10 percent of Europe's shipping container capacity.[35]

Massive PRC investments in both the port infrastructure and port operations of Piraeus, Greece, have drawn attention and consternation by some EU and NATO observers. Piraeus is the only deepwater port in the eastern Mediterranean that has the capacity and infrastructure to allow the transshipment of cargo and the harbor depth to allow the docking of very large container ships. It has a major strategic advantage for container shipping over the northern European ports of Hamburg and Rotterdam because Piraeus's location reduces a week of transit time from Asia at a cost savings of roughly $2 million per trip.[36] Of note to the United States and NATO, Piraeus is also Greece's largest naval base as well as a hub for NATO and the US Navy's Sixth Fleet operations in the Mediterranean Sea.

China recognized the geostrategic value of the port of Piraeus, which Xi Jinping refers to as the "head of the dragon." In 2016, just six years after its initial investment, COSCO became the majority stakeholder in container terminals in Piraeus, operating two of the ports' three terminals via its subsidiary, Piraeus Container Terminal.[37] COSCO also has operational control of the third terminal via its majority stake in Piraeus Port Authority, which increased from 51 to 67 percent following a recent amendment to the 2016 agreement.[38] Piraeus is slated to become the Mediterranean's biggest container port and a major global transshipment node based on the entrepôt model of Singapore.[39]

China's Maritime Silk Road investments in the port of Piraeus have raised alarm in Brussels regarding the susceptibility of financially

35. Joanna Kakissis, "Chinese Firms Now Hold Stakes in over a Dozen European Ports," NPR (website), October 9, 2019, https://www.npr.org/2018/10/09/642587456/chinese-firms-now-hold-stakes-in-over-a-dozen-european-ports#:~:text=In%20the%20past%20decade%2C%20Chinese,of%20Europe's%20shipping%20container%20capacity.

36. John Psaropoulos, "Greece and China Hail Strategic Partnership, as US and EU Look On," *Al Jazeera* (website), November 11, 2019, https://www.aljazeera.com/economy/2019/11/11/greece-and-china-hail-strategic-partnership-as-us-and-eu-look-on.

37. Charlie Lyons Jones and Raphael Veit, *Leaping across the Ocean: The Port Operators behind China's Naval Expansion* (Canberra: Australian Strategic Policy Institute, February 2021), 21, https://s3-ap-southeast-2.amazonaws.com/ad-aspi/2021-02/Leaping%20across%20the%20ocean.pdf?VersionId=mrEJH8QwypEHHxT0jxjtml8ucEeiZJfz.

38. David Glass, "Cosco Completes Increased Stake in Piraeus Port Authority," *SeaTrade Maritime News*, October 12, 2019, https://www.seatrade-maritime.com/ports-logistics/cosco-completes-increased-stake-piraeus-port-authority.

39. Helena Smith, "Xi Jinping Comes to Greeks Bearings Gifts," *Guardian* (website), November 12, 2019, https://www.theguardian.com/world/2019/nov/12/xi-jinping-comes-to-greeks-bearings-gifts.

weaker EU states to foreign influence and coercion. These investments—and other BRI investments discussed later in this chapter—also spurred the need to develop a coherent foreign investment review framework at the EU level, and the enactment of more stringent foreign investment screening measures to monitor acquisitions of strategic critical infrastructure in EU member states.

For NATO, however, there are many more concerning security implications arising from PRC investments in European ports. Of immediate importance is assured access to European port infrastructure and facilities needed to deploy, move, and sustain NATO troops and material quickly and over time in a crisis, conflict, or war. According to one NATO observer, PRC control over European maritime infrastructure "could decrease allies' willingness to move military forces—including sensitive technologies—through the port and its surrounding networks. This could lead to disrupted planning and fewer military exercises, decreasing NATO's ability to defend the Baltic States during a crisis with Russia."[40] Of further note is the reliance on special deepwater facilities required for vessels of the US sealift force. The denial of or impediment to such port access would have a deleterious impact on US force projection capabilities from the homeland as well. Of major concern is that NATO lacks detailed situational awareness of PRC ownership and control of maritime infrastructure and transportation nodes in Europe as well as understanding of how such activity could impede the Alliance's mobility capabilities.

The larger geostrategic issue for NATO regarding the relationship between targeted PRC commercial investments in port infrastructure and control of global port operations is the desire by Beijing to create so-called "strong points" to enable the naval expansion of the People's Liberation Army (PLA) and support future Chinese naval expeditionary warfare capabilities.[41] The PLA navy has enlarged its operations in the Mediterranean and Baltic Seas, and the Arctic region over the past decade. Chinese investments in ports in the Mediterranean Sea, Baltic Sea, Indian Ocean, and along Africa's coastline attend to possible PLA navy basing access, but also offer the potential for sabotage, surveillance, and intelligence collection

40. Lauren Speranza, "China is NATO's New Problem: The Alliance Has Been So Focused on Moscow That It HasMissed Beijing's Growing Clout across Europe," *Foreign Policy* (website), July 8, 2020, https://foreignpolicy.com/2020/07/08/china-nato-hybrid-threats-europe-cyber/.

41. Isaac Kardon, "Research & Debate—Pier Competitor: Testimony on China's Global Ports," *Naval War College Review* 74, no. 1 (Winter 2021): 128, https://digital-commons.usnwc.edu/nwc-review/vol74/iss1/11.

of allied military vessels and operations that routinely use and rely upon these ports.

PRC control over European and Baltic port infrastructure may enable enhanced Sino-Russian military and naval cooperation. Since 2012, China and Russia have held an annual bilateral naval exercise known as "Joint Sea." The intent of this exercise series is to improve tactical and operational capabilities, conduct Joint operations, and increase interoperability between the Chinese and Russian navies. Recent Joint Sea exercises have been held in controversial locations—including the Mediterranean Sea (2015), South China Sea (2016), and the Baltic Sea (2017)—reflecting Beijing's and Moscow's respective support for each other's key security priorities.[42]

These developments have prompted NATO to move mobility and logistics to the forefront of operations regarding force posture and sustainment capabilities. The 2018 NATO Brussels Summit provided a definitive declaration of the importance of force projection, mobility and sustainment: "We are committed to strengthening our ability to deploy and sustain our forces and their equipment, throughout the Alliance and beyond, and aim to improve military mobility by land, air, or sea as soon as possible, but no later than 2024."[43] As a result, two new NATO commands have been established.

First, Joint Force Command (JFC-NF) in Norfolk, Virginia, became operational in 2020. Its mission is to ensure and protect sea lines of communication and security in the Atlantic Ocean, given increasing presence of Russian submarine activity and future PLA Navy activities in the North Atlantic. Second, NATO established the Joint Support and Enabling Command (JSEC) in Ulm, Germany, under the operational command of SACEUR to support, coordinate, and safeguard rapid movement of troops and equipment across European borders.[44] According to NATO spokesperson Oana Lungescu, "The new command in Ulm will help our forces become

42. Alec Blivas, "Sino-Russian Military Exercises Signal a Growing Alliance," *Proceedings* 147, no. 6 (June 2021), https://www.usni.org/magazines/proceedings/2021/june/sino-russian-military-exercises-signal-growing-alliance.

43. "Brussels Summit Declaration," NATO (website), July 11, 2018, https://www.nato.int/cps/en/natohq/official_texts_156624.htm?selectedLocale=uk.

44. Sergei Boeke, "Creating a Secure and Functional Rear Area: NATO's New JSEC Headquarters," *NATO Review*, January 13, 2020, https://www.nato.int/docu/review/articles/2020/01/13/creating-a-secure-and-functional-rear-area-natos-new-jsec-headquarters/index.html.

more mobile and enable rapid reinforcement within the Alliance, ensuring we have the right forces in the right place at the right time."[45]

In addition to creating JFC-NF and JSEC, NATO has significantly expanded its exercises to focus on logistics as a domain of warfare. Exercise Trident Juncture in 2018, for example, was the largest NATO exercise since the Cold War and was intended to showcase the Alliance's ability to respond together in an Article 5 operation.[46] The recent exercise Steadfast Defender 21 tested NATO readiness and military mobility from North America to the Black Sea region, involving both JFC-NF and JSEC. The intent of this exercise—which deployed a division-size force from the United States to Europe, required those forces to pull equipment from Army prepositioned stocks in Europe, and then moved the personnel and equipment across the theater to multiple training areas—was to test the rapid reinforcement by North American forces using a scenario that involved denial of access of critical European ports.[47]

The US European Deterrence Initiative—known as the European Reassurance Initiative until 2018—has been another important mechanism by which the United States has enhanced the forward presence and sustainment of its forces in Europe.[48] The primary focus of the European Deterrence Initiative has been "in resilient joint reception, staging, onward movement, and integration (JRSOI) which has resulted in infrastructure improvements to airfields and other transportation nodes as well as prepositioning of supplies."[49] In concert, the EU—in coordination with NATO—is addressing host nation military logistics through a Permanent Structured Cooperation project focused on enhancing military mobility and ensuring the movement of troops and equipment efficiently across European borders.[50]

45. John Vandiver, "New NATO Command in Germany Will Move Troops and Tanks to Hot Spots," *Stars & Stripes* (website), September, 18, 2019, https://www.stripes.com/theaters/europe/new-nato-command-in-germany-will-move-troops-and-tanks-to-hot-spots-1.599395.

46. Jim Garamone, "NATO Admiral Discusses Complex Security Environment, Results of Trident Juncture," Department of Defense (website), February 21, 2019, https://www.defense.gov/Explore/News/Article/Article/1763179/nato-admiral-discusses-complex-security-environment-results-of-trident-juncture/.

47. "Exercise Steadfast Defender 2021 to Test NATO Readiness and Military Mobility," NATO (website), May 6, 2021, https://www.nato.int/cps/en/natohq/news_183459.htm.

48. See Michelle Shevin-Coetzee, *The European Deterrence Initiative* (Washington, DC: Center for Strategic and Budget Assessments, January 25, 2019), https://csbaonline.org/research/publications/the-european-deterrence-initiative.

49. *United States European Command and United States Transportation Command: Hearing before the Senate Committee on Armed Services*, 116th Cong. (2019) (statement of General Curtis M. Scaparrotti, Commander, US European Command), 15, https://www.armed-services.senate.gov/imo/media/doc/Scaparrotti_03-05-19.pdf.

50. *United States European Command and United States Transportation Command* (2019), 14.

Hybrid Threats to the US and European Defense Industrial Bases

> China, in particular, has made it a national goal to acquire foreign technologies to advance its economy and to modernize its military. . . . It is comprehensively targeting advanced US technologies and the people, the information, businesses and research institutions that underpin them.
>
> —Kari A. Bingen, 2018
> the then US deputy undersecretary of defense for intelligence

To achieve this national goal, China has used an effective combination of industrial, trade and investment policies. China analysts have largely focused on the PRC's illicit means to acquire these technologies through espionage, cyber operations, evasion of US export control restrictions, and through coercive intellectual property sharing requirements on foreign companies investing in the Chinese market. Only recently has the focus turned to Chinese overseas strategic investments—assisted by Beijing-backed investment vehicles, such as the China Investment Corporation, and massive sovereign wealth funds—in the DIB sectors of the United States and NATO member states.[51] Beijing has employed three investment tools: foreign direct investment and acquisitions of US and European dual-use companies; BRI infrastructure investments, particularly in the Baltic states and southern European countries; and the promotion of Chinese state-owned and private sector champions to dominate key markets—especially in telecommunications—as part of Beijing's "Go Out" strategy. Taken together, these activities are potential hybrid threats for the United States and NATO as they have the net effect of the PRC gaining access and control over key DIB infrastructure in a crisis. There is also the longer-term creation of strategic dependencies that enable coercion, intelligence exploitation, and the means to divide Europe and weaken Alliance solidarity.

Regarding Chinese foreign direct investment and acquisitions in the European DIB, examples include the 2008 takeover of the British firm Dynex Semiconductor by a subsidiary of the large state-owned China South Rail. This takeover enabled the PLA to manufacture

51. For additional information, see White House Office of Trade and Manufacturing Policy, *How China's Economic Aggression Threatens the Technologies and Intellectual Property of the United States and the World* (Washington, DC: White House Office of Trade and Manufacturing Policy, June 2018).

insulated-gate bipolar transistor semiconductors—a critical component in electromagnetic aircraft launch systems used for the PLA Navy's next-generation aircraft carriers and railguns.[52] Another high-profile example from 2016 was the PRC Midea Group's acquisition of the robotics firm Kuka, a key player in Germany's industry 4.0.

Beijing has also made significant investments in southern European power grids. For example, the State Grid Corporation of China (SGCC) has invested heavily in grid operators in Portugal and Italy. In 2012, the SGCC became the primary shareholder in Redes Energéticas Nacionais, Portugal's national grid company, making it the dominant shareholder in the Portuguese power grid. Likewise, another Chinese state-owned enterprise, China Three Gorges, has sought to take control by increasing its 23 percent stake in Energias de Portugal, the largest Portuguese power utility company whose assets include the Alqueva Dam. In 2014, the SGCC took a stake in Italy's CDP Reti, which owns gas and power transmission networks, and in 2017, it purchased 24 percent of ADMIE, an independent grid operator in Greece. With these strategic purchases, Beijing controls large interconnections of the southern European grid.

Of note, power grid acquisition strategies of Chinese state-owned actors have not been as successful in northern Europe. For instance, efforts to buy 14 percent of Eandis, a Belgian distributor of gas and electricity, and a 20 percent stake in Germany's high-voltage energy network, 50 hertz, both failed.[53] As the world's premier supplier of transformers, China poses the problem of supply chain dependency and integrity concerns for US and European grid security. In 2020, the then President Donald Trump issued Executive Order 13920, "Securing the United States Bulk-Power System," focusing on the corruption of supply chains for transformers and other bulk power equipment and the danger of adversaries using compromised equipment to cut off the flow of power to defense installations and other critical government facilities.[54]

Both Russia and China have also used foreign investments in strategic geographic locations for intelligence gathering on US and NATO

52. Atlantic Council, *The China Plan: A Transatlantic Blueprint for Strategic Competition* (Washington, DC: Atlantic Council, March 2021), 67, https://www.atlanticcouncil.org/in-depth-research-reports/report/china-plan-transatlantic-blueprint/.

53. "China Eyes Role as World Power Supplier," *Financial Times* (website), June 6, 2018, https://www.ft.com/content/bdc31f94-68aa-11e8-b6eb-4acfcfb08c11.

54. "Executive Order on Securing the Unites States Bulk-Power System," Exec. Order 13920, May 1, 2020, https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-securing-united-states-bulk-power-system/.

military movements. A telling example is the Russian-owned real-estate company in Finland, Airiston Helmi, which was created to house Russian personnel and large transport platforms in a strategically important area of Finland's archipelago. Airiston Helmi is also conveniently located along the major transit strait for cargo vessels, near a basing location for Finnish naval combat vessels, and in the vicinity of key seabed communication cables. Similarly, in 2016, the Anbang Group—a Chinese company with close connections to Beijing—attempted to purchase the iconic Hotel del Coronado on Coronado Island in San Diego, which is home to several key US Navy facilities on the Pacific coast, such as Naval Amphibious Base Coronado, Naval Special Warfare Command, and Naval Air Station North Island, home of Carrier Strike Group 1.[55] Across the Atlantic in Scotland, another Chinese company purchased Rosslea Hall hotel, just a few miles from Faslane, which is the home port to the entire UK fleet of four Vanguard-class ballistic missile nuclear submarines—the sole UK nuclear deterrent—as well as eight nuclear-powered attack submarines.[56]

Recent analyses have focused on deleterious PRC investments in European CI under China's BRI.[57] Under the overarching BRI umbrella, China launched the 17+1 Initiative that includes 12 EU member states and five Balkan countries, with major infrastructure loans going toward the construction of high-speed rail networks, port infrastructure, telecommunications, bridges, and highways. Despite being NATO members, several southern European countries—in particular Italy, Greece, and Portugal—have joined the BRI, providing China with critical gateways and a trans-Eurasian bridgehead to the EU market. Arguably, Chinese BRI investments in Europe are part of a deliberate strategy by Beijing to target the economically weaker EU and NATO members, and draw them into China's orbit. In addition, the PRC's Digital Silk Road undersea telecommunications cables connecting China to Europe have raised the specter of NATO host country communications dependencies on a near peer competitor. Undersea cables have significant strategic importance. Roughly 400 undersea cables carry 98 percent of international Internet data and telephone traffic around the world. US companies have largely owned and operated these telecommunications

55. David Kilcullen, *The Dragons and the Snakes: How the Rest Learned to Fight the West* (New York: Oxford University Press, 2020), 169.

56. Ruth Ingram, "Scotland's Rosslea Hall Hotel and Other CCP Trojan Horses in Britain," Bitter Winter (website), March 17, 2021, https://bitterwinter.org/scotlands-rosslea-hall-hotel-and-other-ccp-trojan-horses-in-britain/.

57. See John R. Deni et al., *Chinese Investment in Post-Pandemic Europe: Security Risks in Infrastructure, Defense Technology, and Political Influence* (Carlisle, PA: Strategic Studies Institute, US Army War College Press, forthcoming).

links, providing a sense of security to the United States and Allies that are concerned about sabotage or surveillance. China's Huawei Marine and Hengtong Optic-Electric Company—of which Huawei Technologies is the third-largest shareholder—are building the highly sensitive "Peace" cable that will travel over land from China to Pakistan, then undersea via the Horn of Africa to its termination in the port city of Marseille, France. Huawei Technologies is also making the equipment for the Peace cable landing stations and its underwater transmission gear. This strategic positioning provides the PRC national telecommunications giant with the ability to divert or monitor data traffic, or, in the event of a conflict, to sever links to nations.[58] One example of this vulnerability in undersea communications is the Arctic Connect data cable, which will link Asia and Europe through the northern sea route along the Arctic coast. Reflecting this global security concern, the then US Secretary of State Mike Pompeo urged the international community to "ensure the undersea cables . . . are not subverted for intelligence gathering by the People's Republic of China at hyper scale."[59]

China's strategic penetration of key DIB infrastructure sectors through the promotion of its state-owned enterprises and national champions—especially in 5G telecommunications—has enormous potential to erode the US and European DIB, and create vulnerable supply chain dependencies. Huawei's recent bid to provide 5G information and communications technology networks in the United States and Europe is highly relevant and poses several security concerns for NATO in part because 5G networks are also far more vulnerable to cyberattacks than their predecessors.[60] For example, should the United States and Europe be dependent on China to provide a key dual-use DIB infrastructure? Through its control of the world's wireless and telecommunications backbone, will the PRC use 5G as a "Trojan horse" for commercial and military espionage and hybrid threat purposes?

Dependency on PRC 5G supply chains raises the additional issue of system and component vulnerabilities. The recent SolarWinds intrusions into US government and private sector networks that were accomplished

58. Atlantic Council, *China Plan*, 66.

59. Bloomberg Businessweek, "China's 7,500-Mile Undersea Cable to Europe Fuels Internet Feud," Bloomberg (website), March 5, 2021, https://www.bloomberg.com/news/articles/2021-03-05/china-s-peace-cable-in-europe-raises-tensions-with-the-u-s.

60. See Tom Wheeler and David Simpson, *Why 5G Requires New Approaches to Cybersecurity: Racing to Protect the Most Important Network of the 21st Century* (Washington, DC: Brookings Institution, September 3, 2019), https://www.brookings.edu/research/why-5g-requires-new-approaches-to-cybersecurity/?amp.

through compromised software supply chains highlight this exact vulnerability. The United States and NATO are highly cognizant of the vulnerability of defense supply chains containing material made in China as well as the consequences of PRC technology embedded in communications systems, especially as most military communications are carried over commercial telecommunications infrastructure. Indeed, the most recent US National Counterintelligence Strategy emphasizes that adversaries are targeting supply chain vulnerabilities and conducting other pre-attack operations so that they can "exploit, disrupt and damage US and Allied critical infrastructure and military capabilities during a crisis."[61]

In the United States, the Trump administration's response to Huawei efforts to dominate the US 5G market was swift and decisive. It banned Huawei from all federal contracts for telecommunications equipment and services, and US government contractors were prohibited from doing business with Huawei as well.[62] The US Department of Justice filed formal charges of fraud, obstruction of justice, and theft of trade secrets against Huawei in January 2019. The Trump administration also exerted considerable political pressure on its allies within the "Five Eyes" intelligence community—comprised of Australia, Canada, New Zealand, the United Kingdom, and the United States—to ban Huawei from their respective markets.

Concerned about the larger implications of Chinese investments and other adversarial activities involving the DIB infrastructure, the US Congress passed the Foreign Investment Risk Review Modernization Act of 2018, which expanded the powers of the Committee on Foreign Investment in the United States to prevent foreign adversaries from gaining control of defense industrial infrastructure assets.[63] In 2018, the then president Trump issued Executive Order 13806, which mandated an assessment of the broader US DIB sector. That assessment concluded, "All facets of the manufacturing and defense industrial base are currently under threat, at a time when strategic competitors and revisionist powers appear to be growing in strength and capability."[64]

By contrast, Europe has been divided in recognizing and acknowledging the potential security vulnerabilities and dependencies created by Chinese

61. National Counterintelligence Security Center, *National Counterintelligence Strategy: 2020–2022* (Washington, DC: White House, January 7, 2020), 3.

62. National Defense Authorization Act of 2019, P.L. 115-232 (2018), sec. 889(f)(3).

63. National Defense Authorization Act of 2019, sec. 1701-1703.

64. Department of Defense (DoD), *Assessing and Strengthening the Manufacturing and Defense Industrial Base and Supply Chain Resiliency of the United States* (Washington, DC: DoD, September 2018), 8.

investments in infrastructure. Even as Europe is awakening to the strategic implications of PRC investments, the approach to China from governments within the Alliance remains very uneven. In October 2019, the then chancellor Angela Merkel allowed China's Huawei and ZTE greater market access into Germany's 5G networks. Since Germany is both a key NATO Ally and China's largest bilateral trading partner, Merkel's decision has had multiple international security implications. One NATO analyst contends that it threatens NATO security and the operations of the US armed forces based in Germany, and contravenes US intelligence warnings.[65]

By leveraging its presence in Europe's DIB infrastructure, Beijing has wielded powerful political influence and effective veto power within the EU on several occasions. Hungary and Greece, for example, sought to block any direct reference to China in an EU statement regarding the ruling by the Permanent Court of Arbitration in The Hague that struck down the PRC's legal claims in the South China Sea.[66] In another high-profile incident, Greece blocked an EU statement at the United Nations criticizing China's human rights record.[67] Sounding the alarm over the long-term implications of European BRI investments on EU unity, Sigmar Gabriel—the then foreign minister of Germany—warned, "If we do not succeed, for example, in developing a single strategy toward China, then China will succeed in dividing Europe."[68]

The security and dependency of host nation critical infrastructure—such as data cables, 5G networks, electricity grids, transportation, and logistics infrastructures—on China has alarmed NATO. Chinese involvement in key infrastructure projects in Europe has raised concern and garnered increasing attention by NATO regarding Beijing's intentions and the need for a shared Allied policy on China. On the occasion of NATO's 70th anniversary meeting in London in December 2019, Secretary General Jens Stoltenberg warned, "What we see is that the rising power of China

65. John R. Deni, "Opinion: Germany's Refusal to Ban China's Huawei from 5G Is Dangerous for the West," *Newsweek* (website), October 30, 2019, https://www.newsweek.com/germanys-refusal-ban-chinas-huawei-5g-dangerous-west-opinion-1468520.

66. Erik Brattberg and Etienne Soula, *Europe's Emerging Approach to China's Belt and Road Initiative* (Washington, DC: Carnegie Endowment for International Peace, October 18, 2019), https://carnegieendowment.org/2018/10/19/europe-s-emerging-approach-to-china-s-belt-and-road-initiative-pub-77536.

67. Robin Emmott and Angeliki Koutantou, "Greece Blocks EU Statement on China Human Rights at UN," Reuters (website), June 18, 2017, https://www.reuters.com/article/us-eu-un-rights/greece-blocks-eu-statement-on-china-human-rights-at-u-n-idUSKBN1990FP.

68. Lucrezia Poggetti, "China—One Europe? German Foreign Minister's Remarks Irk Beijing," *Diplomat* (website), September 9, 2017, https://thediplomat.com/2017/09/one-china-one-europe-german-foreign-ministers-remarks-irk-beijing/.

is shifting the global balance of power . . . we have to address the fact that China is coming closer to us, investing heavily in infrastructure. . . . So, of course, this has some consequences for NATO."[69]

A recent NATO report was more direct in identifying the potential consequences that adversaries' ability to penetrate DIB infrastructure would have on NATO security: "The degree and impact of foreign direct investment in strategic sectors—such as airports, sea ports, energy production and distribution, or telecoms—in some Allied nations raises questions about whether access and control over such infrastructure can be maintained, particularly in crisis when it would be required to support the military."[70] Incremental progress has been made recently with a new EU regulation establishing a framework for screening foreign direct investments in European critical infrastructure and technologies. In 2019, the EU established a foreign investment review framework on the grounds of security or public order but left individual member states with the authority for screening and decision making.[71] As with issues of energy security, NATO is grappling with dependency on European host country infrastructure and the vulnerabilities this dependency poses for logistics, secure communications, interoperability, and other requirements to enable mobilization, force projection, and sustainment.

NATO Measures to Redress Vulnerabilities from Hybrid Threats

NATO has developed a multifaceted and loosely coordinated approach to overcome the critical infrastructure vulnerabilities created by hybrid threats from adversaries. Initial Allied efforts have focused on building organizational capacity with the establishment of NATO Centres of Excellence (COE) to support critical infrastructure protection. Key examples include the COEs for: Defence Against Terrorism (Ankara, Turkey) in 2004; Cooperative Cyber Defence (Tallinn, Estonia) in 2008; Energy Security (Vilnius, Lithuania) in 2012; and Maritime Security (Istanbul, Turkey) in 2020. The activities

69. Holly Ellyatt, "China is 'Coming Closer' but We Don't Want a New Adversary, NATO Chief Says," *CNBC* (website), December 2, 2019, https://www.cnbc.com/2019/12/02/jens-stoltenberg-rising-power-china-must-be-addressed-by-nato.html.

70. Wolf-Diether Roepke and Hasit Thankey, "Resilience: The First Line of Defense," *NATO Review* (website), February 27, 2019, https://www.nato.int/docu/review/articles/2019/02/27/resilience-the-first-line-of-defence/index.html.

71. *Regulation (EU) 2019/452 of the European Parliament and of the Council of 19 March 2019 Establishing a Framework for the screening of Foreign Direct Investments into the Union*, European Parliament, https://eur-lex.europa.eu/eli/reg/2019/452/oj.

of these centers have been bolstered by the work of the European COE for Countering Hybrid Threats based in Helsinki, Finland. Together, these COEs raise awareness and support the security and resilience of critical infrastructure through training, education, and sharing best practices among government, academic, private sector, and other subject-matter experts.

In parallel, NATO has developed a much-needed policy framework for how to respond to hybrid threats to NATO's collective defense. This policy was promulgated at the 2018 Brussels Summit. Per the summit declaration, NATO affirmed that while "the primary responsibility for responding to hybrid threats rests with the targeted nation, NATO is ready, upon Council decision, to assist an Ally at any stage of a hybrid campaign. In cases of hybrid warfare, the Council could decide to involve Article 5 of the Washington Treaty, as in the case of armed attack."[72]

NATO has also developed and deployed tools to strengthen critical infrastructure resilience among member states. NATO's hybrid threat toolbox includes greater situational awareness through more deliberate intelligence and information sharing via the creation of a joint intelligence and security division at NATO headquarters in 2017, as well as other information sharing mechanisms with the EU and the private sector. See chapter 11 for additional best practices in information and intelligence sharing. The Cooperative Cyber Defence COE has enhanced assistance for cyber defense through the deployment of mobile computer emergency response teams and better coordination of computer emergency and incident response capabilities between the EU and NATO. See chapter 14 for an overview of recommendations for improving cybersecurity capabilities.

Other initiatives have focused on analyzing NATO member states' energy interdependencies and vulnerability to hybrid threats and interference, as well as mitigation measures to increase the resilience of energy infrastructure in Europe. See chapter 5 for discussion of these threats and of integrating Ukraine's energy infrastructure into the European grid. NATO is increasingly using exercises and war games—such as Defender-21 and Locked Shields—to identify hybrid threats to critical infrastructure that would undermine NATO force deployment and logistics capabilities. Locked Shields, for instance, is an annual exercise that enables cybersecurity experts to enhance their skills in defending national information technology systems and other critical infrastructure under real-time attacks. Additionally, NATO has established counter-hybrid support teams to be utilized in a crisis or to assist

72. "Brussels Summit Declaration."

in building national counter-hybrid capacities. NATO first deployed one of these counter-hybrid teams in November 2019 to assist Montenegro, then a new NATO Ally, in its efforts to mitigate critical infrastructure hybrid threat vulnerabilities.[73]

Recognizing that NATO is dependent on host country critical infrastructure that is owned and operated largely by the private sector, recent NATO efforts have been concentrated on improving the resilience of Allied civilian infrastructure to respond to and recover from hybrid threats and attacks. Private-sector critical infrastructure owners and operators, driven by business models of cost, often have not invested in building redundancies or implementing cybersecurity measures, thereby making them vulnerable to hybrid threats—as a plethora of recent cyberattacks against key global critical infrastructure illustrates all too well.

The principle of resilience is established in Article 3 of the Washington Treaty, which identifies NATO's first line of defense as the responsibility of each Ally to "maintain and develop individual and collective capacity to resist armed attack" through "continuous and effective self-help and mutual aid."[74] NATO recognizes that a higher level of national resilience is increasingly required in the dynamic hybrid threat environment as outlined above. At the 2016 Warsaw Summit, NATO announced the implementation of seven baseline country requirements for measuring and improving national resilience and civil preparedness. These requirements involve strengthening resilience in seven strategic sectors: continuity of government, energy, population movement, food and water resources, mass casualties, civil communications, and transport systems.[75] To support this effort, NATO created resilience advisory support teams to provide expertise to help member states assess and build resilience.

These various measures have been tremendously important in bringing the twin issues of security and resilience of critical infrastructure to the forefront and engendering much-needed coordination between and among the EU, NATO, and individual states. NATO, however, still needs to decipher key host country infrastructure dependencies and identify where vulnerabilities

73. Michael Rühle and Clare Roberts, "Enlarging NATO's Toolbox to Counter Hybrid Threats," *NATO Review* (website), March 19, 2021, https://www.nato.int/docu/review/articles/2021/03/19/enlarging-natos-toolbox-to-counter-hybrid-threats/index.html.

74. "Resilience and Article 3," NATO (website), June 25, 2018, https://www.nato.int/cps/en/natohq/topics_132722htm.

75. Wolf-Diether Roepke and Hasit Thankey, "Resilience: The First Line of Defense," *NATO Review* (website), February 27, 2019, https://www.nato.int/docu/review/articles/2019/02/27/resilience-the-first-line-of-defence/index.html.

in private or foreign ownership of critical national infrastructure could impede its missions related to force projection, mobility, and sustainment operations in a contested hybrid threat environment. Identifying key mission functions to supporting critical infrastructure and their associated vulnerabilities is the approach that the department has implemented in the United States and that could be expanded to include NATO.

With the introduction of the Mission Assurance Strategy in 2012, there was a major paradigm shift from protecting defense CI assets toward strengthening the resilience of DoD missions. The strategy defines mission assurance as a "process to protect or ensure the continued function and resilience of capabilities and assets—including personnel, equipment, facilities, networks, information and information systems, infrastructure, and supply chains—critical to the performance of DoD MEFS [mission essential functions] in any operating environment or condition."[76] Recognizing that over 90 percent of US infrastructure resides in the private sector, the strategy also calls for strengthening DoD partnerships with those commercial infrastructure owners and operators. The strategy has been augmented by other policy directives that require and provide all DoD services, departments, and agencies with guidelines and procedures for identifying, assessing, managing, and monitoring risks to strategic missions.[77] At present, mission assurance programs at the levels of the Joint Staff, the Office of the Secretary of Defense, and the Services apply to installations located in the continental United States and the supporting domestic critical infrastructure. Hence, there is a key opportunity for the DoD to collaborate with NATO to implement a mission assurance-based program that would provide a rigorous risk assessment and management framework, model and identify critical infrastructure dependencies using subject-matter expertise—such as Argonne National Laboratory—and provide well-proven mitigation measures to ensure critical infrastructure resilience and redundancy of vital nodes for NATO mission sets.

76. Office of the Undersecretary of Defense (Policy) [OUSD (P)], *Mission Assurance Strategy* (Washington, DC: DoD, April 12, 2012), 1, https://policy.defense.gov/Portals/11/Documents/MA_Strategy_Final_7May12.pdf.

77. See OUSD (P), *Mission Assurance*, DoD Directive 3020.40 (Washington, DC: DoD, November 29, 2016); and OUSD (P), *Mission Assurance (MA) Construct*, DoD Instruction 3020.45 (Washington, DC: DoD, August 14, 2018).

Conclusion

Adversaries are actively targeting US and NATO critical infrastructure to undermine military capability, force projection, mobility, and sustainment. In some cases, adversaries are penetrating this infrastructure to identify vulnerabilities for later exploitation, and in others, critical infrastructure is being weaponized as a form of hybrid threat. While the United States and NATO grapple with the challenges of this ongoing and future form of threat activity, partner countries are well-advised to learn from their experiences. Internationalization of investments and supply chains—coupled with rapid advances in communications, computing, artificial intelligence and other technologies—are creating new global and regional critical infrastructure interdependencies. These new infrastructure interdependencies will place a premium on greater civil-military and public-private sector coordination.

— 5 —

European Energy and the Case of Ukraine

Theresa Sabonis-Helf

Developed states require a continuous influx of energy resources from either domestic or foreign sources. In wartime, the ability to supply fuel to the battlefield can be decisive. Even in peacetime, denial of energy resources can be catastrophic. Interruption of oil supply can cause transportation to grind to a halt; interruption of gas supply plunges societies into heating, electricity, and industrial crises; and interruption of electricity casts populations into darkness and silence. Although any supply disruption poses critical challenges, the vulnerability of electricity systems is increasingly urgent. Both the data intensity of everyday life and the societal effort to reduce greenhouse gas emissions are moving modern societies increasingly toward electrification.

While electrification is more important in the twenty-first century, the West's electricity infrastructure remains a product of the twentieth century, posing new strategic challenges to collective thinking about security and resilience. Potential electricity interruption is becoming both increasingly catastrophic for urbanized areas and more attractive to actors seeking disruption. Simultaneously, the avenues for disruption are becoming greater; as systems become larger, "smarter," and more internationally linked, the new

technologies being incorporated into the grid invite new avenues of attack.[1] The intertwined relationship between electricity security and cybersecurity calls for an understanding of critical infrastructure security and resilience (CISR) that recognizes both sets of vulnerabilities. Here, the case of Ukraine is instructive. Ukraine's experience of energy security and cybersecurity in recent years reveals significant risks and offers insight into the North Atlantic Treaty Organization's efforts to enhance civil preparedness and collective CISR among Allies and partners, including Ukraine. It also illustrates the complexities ahead.

The shift in demand of all energy toward electrical energy is evident in market and policy trends. In 2018, electricity accounted for 19 percent of total world energy consumption, but that number is rising rapidly. In its 2021 *Net Zero by 2050* report, the International Energy Agency (IEA) noted that, if the world is to meet its climate change abatement goals, electricity should comprise nearly half of all energy by 2040.[2] This was a significant upward revision from its 2019 prediction that electricity would increase to between 20–31 percent by that date.[3] Based on its stated policies, the European Union is expected to lead this transition. EU policies favor investing in cross-border electricity projects, and policies aiming for at least 30 million electric vehicles by 2030 have caused electric vehicle sales to triple since 2019.[4]

Substantial gains in efficiency, room for more renewable energy, and emissions reductions can all result from growing reliance on electricity grids, but this change also represents a shift in critical infrastructure risk. Drinking and wastewater systems, food, health care, communications, and financial services depend on reliable electricity. See chapter 1 for discussion of the relationship between lifeline sectors and other critical infrastructure sectors. In addition, the world faces rising demand for electricity from data centers, which already rank among the largest consumers of electricity.

Acknowledgments: The author would like to express her gratitude to Rayanne Fujimoto and Hans Johnson, outstanding graduate students in the Georgetown University Master of Science in Foreign Service Program, who made substantial contributions to this chapter.

1. National Commission on Grid Resilience (NCGR), *Grid Resilience: Priorities for the Next Administration* (Washington, DC: NCGR, 2020), 5.

2. International Energy Agency (IEA), *Net Zero by 2050* (Paris: IEA, 2021), 27, https://www.iea.org/reports/net-zero-by-2050.

3. IEA, World Energy Outlook 2019 (Paris: IEA, 2019), 253, https://www.iea.org/reports/world-energy-outlook-2019/electricity.

4. Kate Abnett, "Electric Car Sales Surge as Europe's Climate Targets Bite," Reuters (website), June 29, 2021, https://www.reuters.com/business/sustainable-business/electric-car-sales-surge-europes-climate-targets-bite-2021-06-29/.

Experts estimate the global electricity demand of data centers will increase 15-fold by 2030.[5] Further electrification of the fleet will make transport more dependent as well, and integration of grids increases the ability to spread vulnerabilities.

Although this shift toward electrification brings new risk, energy-importing states have demonstrated the ability in the past to meet new energy sector risks successfully. This chapter will offer a brief overview of the evolution of European concerns and approaches to securing energy supply, with an emphasis on the contemporary issues associated with securing power grids. Ukraine has suffered dramatic power grid disruptions in recent years, and Europe is pursuing a strategy of enhancing Ukraine's energy security by tying its systems more closely to Europe—a strategy that necessitates significant shifts in policy and strategy for both Ukraine and the EU.

Brief History of European Energy Security Concerns

The security of energy supply has long been a strategic preoccupation of importing states. When Winston Churchill, the then first lord of the admiralty, announced his decision in 1911 to convert the Royal Navy from coal to oil battleships, he made a conscious decision to trade the security of domestic coal supply for the speed and flexibility of oil, even though it would be imported. He famously argued that "safety and certainty in oil lie in variety and variety alone."[6] Secure flow of supply—through multiple routes, numerous suppliers, and secure transport—has remained a military and foreign policy priority for European states ever since. In each era, energy supply crises sparked innovation and strategies for collective security. In World War II, Operation Pluto, a clandestine project to create and operate a cross-Channel undersea pipeline supplying petroleum, was critical to the success of the Normandy invasion, and set the precedent for undersea pipelines that are now a staple of energy transmission.[7] Energy supply was a vulnerability for both

5. Kate Crawford, *Atlas of AI* (New Haven, CT: Yale University Press, 2021), 42–43.

6. Daniel Yergin, *The Quest: Energy, Security, and the Remaking of the Modern World* (New York: Penguin Books, 2011), 265.

7. Arnold Krammer, "Operation PLUTO: A Wartime Partnership for Petroleum," *Technology and Culture* 33, no. 3 (1992): 441–42, https://doi.org/10.2307/3106633.

sides, as historians offer ample evidence that the Axis powers both suffered from key vulnerabilities in oil and sought to disrupt Allied supply.[8]

The first great shock to supply in peacetime was the 1973 oil crisis. The panic and price spiral that resulted from targeted reductions in supply led importing states to establish the IEA. Under the IEA treaty, members are obliged to create strategic reserves of oil, equal to 90 days of net oil imports, and share the reserves in crisis.[9] Although mandatory sharing of the reserves has been invoked only three times in the IEA's history, maintaining strategic reserves has come to be regarded as a best practice.

In the wake of the Iranian Revolution, the world experienced a second oil shock in the 1970s, which led Europe to begin considering alternative fuels and embrace imported natural gas. With West Germany in the lead, Europe determined to build a natural gas pipeline from Siberia to Europe as a solution to the problem of overreliance on an unstable Persian Gulf. Despite US objections and equipment embargoes, the 3,300-mile pipeline with a capacity of 32 billion cubic meters per year was completed in 1984.[10] Europe pledged that it would not become more than 30 percent dependent on Soviet gas supply. This "acceptable level of dependence" became standard practice during the Soviet period and incentivized pursuit of gas pipeline imports from North Africa to supplement Soviet natural gas. According to energy scholar Thane Gustafson, economic necessity overcame political concerns for both the Soviets and Europeans. The Soviet Union, whose economy by the 1980s was collapsing from decades of centralized control, became as dependent on hard currency revenues from natural gas sales as Europe was on Soviet hydrocarbons.[11] Moscow, aspiring to keep this essential market, safeguarded a reputation for reliability and proved to be a consistent, non-politicized supplier to states that paid world prices.

Europe enjoyed the advantages of natural gas, which creates less local pollution and greenhouse gases compared to other fossil fuels. It became popular for electricity as well as heat and industry, providing more flexibility

8. Daniel Yergin, "Blood and Oil: Why Japan Attacked Pearl Harbor," *Washington Post* (website), December 1, 1991, https://www.washingtonpost.com/archive/opinions/1991/12/01/blood-and-oil-why-japan-attacked-pearl/1238a2e3-6055-4d73-817d-baf67d3a9db8/; and Marshall I. Goldman, *Petrostate: Putin, Power, and the New Russia* (New York: Oxford University Press, 2008), 33–54.

9. "Oil Stocks of IEA Countries," International Energy Agency (website), June 11, 2021, https://www.iea.org/articles/oil-stocks-of-iea-countries.

10. Meghan L. O'Sullivan, *Windfall* (New York: Simon & Schuster, 2018), 167–68.

11. For additional information, see Thane Gustafson, *Crisis Amid Plenty: The Politics of Soviet Energy under Brezhnev and Gorbachev* (Princeton, NJ: Princeton University Press, 1989).

for power producers. Unlike coal and nuclear power plants, electricity output from gas plants can be easily calibrated to meet fluctuating demand. Natural gas-fired power plants can therefore operate at a profit and with less financial risk across a wide range of demand scenarios, while sources such as coal or nuclear lose money if not operated close to full capacity. Gas plants also require less up-front capital investment and environmental permitting than coal, nuclear, and hydro plants, and thus have much shorter lead times.

Given these advantages, it is not surprising natural gas began to dominate electricity production in Europe and was increasingly used in industry as well. Gas posed a new challenge, however, in terms of security. While oil supply could be secured by multiple routes, natural gas supply requires constant load and pressure, is more difficult to store than oil or coal, and—because it has to be managed across the supply chain—gives power to the transmission system operators.[12] Energy scholar Margarita Balmaceda terms the problem *networkness*, or the "degree to which the overall functioning of the system may be dependent on the network working properly *as a network*."[13]

The degree to which Europe relied on a source that was highly network-dependent became evident with the 2004 expansion of the EU to include states that already had Soviet infrastructure tying them to Russia, and long-standing bilateral energy relationships with Moscow. By contrast to the European experience of non-politicized supply, the former Eastern bloc states had long experience with politicization of both price and supply. The Czech Republic, all three Baltic states, Hungary, Poland, Slovakia, and Slovenia each had long-standing energy relationships with Moscow, and Russia attempted to continue its bilateral, by-country approach, reinforced by annual contracts and old Soviet pipeline infrastructure. As time went on, and tensions rose with specific states, Russia found all of Europe getting involved in what it had previously seen as bilateral agreements. In addition, transit problems began to emerge. Territory that had been part of one former Soviet state now became multiple sovereign states, and the logistics and logic of transit became much more complex as each state pursued its own interests and faced its own transition challenges.

Russian disputes with Ukraine in the winters of 2005, 2006, and 2007—stemming from Russia's disapproval of anti-Russian political actors, but also from Ukraine's failure to pay for its contracted gas imports and disappearance

12. Margarita M. Balmaceda, *Russian Energy Chains: The Remaking of Technopolitics from Siberia to Ukraine to the European Union* (New York: Columbia University Press, 2021), 210.

13. Balmaceda, *Russian Energy Chains*, 64.

of gas from the strategic reserves—came to involve Europe. In December 2008, when bilateral negotiations broke down and Russia exited the talks vowing to reduce flows of gas to Ukraine, the ensuing crisis took 21 days to resolve. Murmurs of concern about energy vulnerability grew louder. By the end of this crisis, Europe resolved to find an approach that would limit Russia's power in natural gas supply.[14] NATO also responded, declaring a role for the Alliance in energy security at the Bucharest Summit in 2008 and establishing the NATO Energy Security Centre of Excellence in Vilnius in 2012.[15]

In 2009, Europe undertook significant changes to laws governing energy trade between Russia and Europe. Recognizing the challenge of *networkness*, the EU adopted legislation known as the Third Energy Package to connect European energy markets in new ways and reduce Moscow's ability to pressure a single member. In addition, the EU designated gas and electricity as priority areas for active intervention in infrastructure development.[16] At the time, many European states lacked interconnectors which would enable them to supply each other with gas in a crisis. Establishing such an interconnector system became a critical component of limiting Moscow's ability to shut off natural gas to select countries. Prompted by the first European directive on security of gas supply, inter-European interconnection capacity significantly improved—increasing 18 percent between 2009 and 2017—as Europe invested in four new cross-border pipelines and converted nine existing lines to two-directional interconnections during this period.[17] Further improvements were made by requiring transit states to become gas hubs instead, equipped to import gas from at least three sources, and possessing natural gas storage capability.[18] These two requirements led to a significant increase in liquefied natural gas (LNG) import capability, and thereby reduced the threat to supply from overland monopoly sources.

Actual physical destruction of pipelines did not touch Europe in this era, but in the former Soviet Union several infrastructure crises led to the

14. Theresa Sabonis-Helf, "Russia and Energy Markets," in *New Realities: Energy Security in the 2010s and Implications for the US Military* (Carlisle, PA: Strategic Studies Institute, US Army War College Press, February 2015), 15–45.

15. Arnold C. Dupuy et al., "Energy Security in the Era of Hybrid Warfare," *NATO Review* (website), January 13, 2021, https://www.nato.int/docu/review/articles/2021/01/13/energy-security-in-the-era-of-hybrid-warfare/index.html.

16. Balmaceda, *Russian Energy Chains* , 212.

17. Yassine Rqiq et al., "Assessing the Impact of Investments in Cross-Border Pipelines on the Security of Gas Supply in the EU," *Energies* 13, no. 11 (2020): 11, https://doi.org/10.3390/en13112913.

18. "Bulgarian Gas Hub Plans Raise Eyebrows in Brussels," *Euractiv* (website), June 27, 2018, https://www.euractiv.com/section/energy/news/bulgarian-gas-hub-plans-raise-eyebrows-in-brussels/.

creation of alternate routes. The Chechens in Grozny repeatedly destroyed the pipelines bringing oil from Azerbaijan into Russia, which caused Azerbaijan to begin reconstruction in 1994 of a disused pipeline, the Baku-Supsa oil pipeline. Commissioned in 1999, that modest-capacity line (145,000 barrels per day) which transited Georgia to the Black Sea provided proof of concept that Georgia could serve as an effective transit state despite its civil conflicts. Construction of the Baku-Tbilisi-Ceyhan oil pipeline and the parallel Baku-Tbilisi-Erzurum natural gas pipeline followed, both coming online in 2006 despite Russian pressure.

Europe also met the security of supply threats with diversification of pipeline routes. New lines including Nordstream I, the recently completed Nordstream II, Turk Stream, and the Southern Gas Corridor—although only the last carries non-Russian gas—help reduce the threat of interruption at a single site. Europe has further ensured security of supply with the increased use of LNG. The significant cost difference between pipeline gas and LNG, however, makes the latter more attractive as an emergency source than a continuous one. Prior to the COVID-19 pandemic, Europe's average utilization rate of its existing LNG terminals was just 25 percent.[19] In addition, the market is changing. The European Investment Bank stated its intent to phase out investment in gas infrastructure by the end of 2022.[20] It may be optimistic to assume that the net-zero goals will be met, since the bank has already changed its original 2021 deadline for phasing out fossil fuels, but the cost-effectiveness of redundant natural gas infrastructure is very much in question.[21]

As the previous examples have shown, a change in source—from coal to oil or oil to gas—shifts the risk as well. The new age is no exception. The shift toward renewable energy gives a growing role to electricity, and if Europe is on the verge of an electricity age, the twentieth-century grid must be better understood and secured against twenty-first century threats. Renewable energy is generally produced domestically, reducing dependence on the long import chains characteristic of fossil fuels. But in order to make grids work most effectively, the engineering imperative is to (1) connect

19. King & Spalding, *LNG in Europe 2018: An Overview of LNG Import Terminals in Europe* (Atlanta: King & Spalding, 2018), 29, https://www.kslaw.com/attachments/000/006/010/original/LNG_in_Europe_2018_-_An_Overview_of_LNG_Import_Terminals_in_Europe.pdf?1530031152.

20. Kira Taylor, "Not Quite over Yet: EIB Spent Euro 890 Million on Fossil Gas since Phase Out, Activists Say," *Euractiv* (website), May 4, 2021, https://www.euractiv.com/section/energy-environment/news/not-quite-over-yet-eib-spent-e890-million-on-fossil-gas-during-phase-out-activists-say/.

21. Jonas Ekblom, "European Investment Bank to Cease Funding Fossil Fuel Projects by End-2021," Reuters (website), November 14, 2019, https://www.reuters.com/article/us-climate-europe-eib/european-investment-bank-to-cease-funding-fossil-fuel-projects-by-end-2021-idUSKBN1XO2OS.

them across time zones so that fluctuations in demand are less severe, (2) include countries with differing sources so that seasonal hydro or time-of-day solar power is optimized, and (3) establish highly flexible markets so that electricity purchases can shift readily. This electricity grid, from a critical infrastructure perspective, has even more *networkness* than natural gas supply lines. A shared grid is one in which domestic electricity must be secured against outside attack, and energy partners, even if they are allies, must be held to high standards so that their system vulnerabilities do not radiate out to the larger grid.

Europe is determined to enter the new era of electricity. Under the European Climate Law adopted in 2020, the EU is required to cut greenhouse gas emissions by 55 percent below the 1990 baseline by 2030. Experts estimate this means that fossil fuel consumption would need to decline some 36 percent from 2020 to 2030, with electricity filling the gap.[22] The shift toward electricity transmission and storage is clear in EU policy and funding priorities. The 2019 list of EU Projects of Common Interest, traditionally focused on strategic energy projects, includes 149 projects, of which 100 are focused on electricity transmission and storage, and six are smart grid deployments.[23] The EU Directorate-General for Energy develops and implements a European energy policy aimed at establishing a competitive and affordable market for technologically advanced energy; promoting sustainable energy production, transportation, and consumption in line with the EU's 2050 decarbonization goals; and enhancing conditions for safe and secure energy supply "in a spirit of solidarity between EU countries ensuring a high degree of protection for European citizens."[24] These desires led the EU to emphasize the importance of extending the European grid to neighboring states even as Europe strives to improve the resilience of the existing grid.

Shifting markets and proliferating routes have not assuaged NATO's concern about the rise in threats to energy, and the new threat in an era of transboundary electricity is an emerging strategic challenge. NATO defined three roles for itself in energy security: (1) raising awareness,

22. Mason Inman, Greig Aitken, and Scott Zimmerman, *Europe Gas Tracker Report 2021* (San Francisco: Global Energy Monitor, April 2021), 8, https://globalenergymonitor.org/wp-content/uploads/2021/03/GEM-Europe-Gas-Tracker-Report-2021.pdf.

23. "Key Cross Border Infrastructure Projects," European Commission (website), July 2, 2020, https://ec.europa.eu/energy/topics/infrastructure/projects-common-interest/key-cross-border-infrastructure-projects_en.

24. "European Commission Directorate General for Energy Explores World Energy Council Global Scenarios," World Energy Council (website), March 13, 2017, https://www.worldenergy.org/news-views/entry/european-commission-directorate-general-for-energy-explores-world-energy-council-global-scenarios.

to include intelligence-sharing on energy development; (2) supporting protection of critical energy infrastructure through the inclusion of energy-related scenarios in exercises, sharing best practices, and providing training courses; and (3) enhancing energy efficiency in the military.[25] While energy remains a largely domestic and nonmilitary issue, NATO has increased its attention to the problems, and has also recognized the important connection between hybrid conflict and energy. See chapter 4 for more detail on the nature of hybrid threats.

Beginning with the Chechen wars, Russia pursued a hybrid strategy of conflict in former Soviet regions. Since hybrid attacks seek low-cost, high-yield ways to influence the policies of states, NATO has recognized that energy is an attractive target. In 2020, the NATO Science and Technology Board formally created a research task group dedicated to energy security in an era of hybrid warfare. This group has the task of analyzing "the hybrid-energy threat and its impact on NATO's military preparedness and ability to execute a mission, its members' infrastructural resilience and ability to participate in a NATO mission, and, ultimately, the coherence of the Alliance."[26] The case of Ukraine is likely to feature prominently in their assessment, since it provides a rich array of examples of the links between hybrid warfare and twenty-first-century energy security.

The Case of Ukraine

Ukraine offers multiple examples of the troubling intersection of energy security, hybrid warfare, and cyberattacks. Russian actors successfully attacked the Ukrainian grid in December 2015, in what is recognized as the first hacker-induced blackout of a region and among the first cyberattacks to cause physical disruption of systems. Cyber techniques brought down Ukraine's grid again in December 2016. The second attack was more sophisticated and designed (albeit unsuccessfully) to do longer-term damage to the grid. A year later, yet another cyberattack—targeted at Ukraine's business sector rather than its infrastructure—damaged the electricity grid yet again, but it more famously created an estimated $10 billion in damage, most of which was outside of Ukraine. Although they differed in targets and sophistication, each cyberattack had some common elements and, most importantly, a common origin. The US Department of Justice found

25. Julijus Grubliauskas and Michael Rühle, "Energy Security: A Critical Concern for Allies and Partners," *NATO Review* (website), July 26, 2018, https://www.nato.int/docu/review/articles/2018/07/26/energy-security-a-critical-concern-for-allies-and-partners/index.html.

26. Dupuy et al., "Energy Security."

in 2020 that the Sandworm group, sponsored by Russia's Main Intelligence Directorate, was behind all three attacks.[27]

NATO countries contributed significantly to getting the lights back on as rapidly as possible, and to reducing future vulnerabilities. They found, however, that identifying the source of the cyberattack took some time, thus making it difficult for supporters to organize a diplomatic response. The cyberattacks were particularly problematic for the EU, which had already made commitments to integrate Ukraine into its own grid.

Setting the Ukrainian Context: Early Energy Conflict

In spite of its shock value, the first attack on Ukraine's grid in December 2015 was not a "bolt out of the blue." Energy had already played an important role in Russia's war objectives in Ukraine. Extensive development underway in the Black Sea off the coast of Crimea would have reduced Ukraine's energy dependence on Russia and thereby Russia's leverage. Russian victory in Crimea cost Ukraine some 80 percent of its oil and gas deposits in the Black Sea due to uncertainty about offshore rights.[28] In addition, since the major Soviet legacy natural gas trunk lines lie below the Donbas, Russia assumed that physical control of these lines would likely increase Russia's control of Ukraine's behavior. Energy is therefore rightly regarded as playing a role in Russian objectives in the conflict.

Ukraine's behavior as a transit state—especially its theft of Russian gas from the large strategic reserves—contributed to declining reliability of supply for Europe and increasing concern on Europe's part regarding its dependence on Russia. Since 2014, Ukraine has placed its gas storage facility—one of the world's largest—under European regulatory structures, thus increasing transparency, destroying Russia's ambition to take over regulation of the facility, and providing added security to Europe against Russian price and supply manipulation. As a reflection of the energy-related tensions between the states, Ukraine stopped purchasing natural gas directly from Russia in 2015 and now reimports gas from Europe instead, even though it remains a transit state for Russia. EU-US cooperation ensured that Russia would contract

27. Department of Justice, "Six Russian GRU Officers Charged in Connection with Worldwide Deployment of Destructive Malware and Other Disruptive Actions in Cyberspace," Department of Justice Office of Public Affairs (website), October 19, 2020, https://www.justice.gov/opa/pr/six-russian-gru-officers-charged-connection-worldwide-deployment-destructive-malware-and.

28. Ariel Cohen, "As Russia Closes in on Crimea's Energy Resources, What Is Next for Ukraine?," *Forbes* (website), February 28, 2019, https://www.forbes.com/sites/arielcohen/2019/02/28/as-russia-closes-in-on-crimeas-energy-resources-what-is-next-for-ukraine/?sh=5495736f29cd.

to use Ukraine as a transit state until 2024, despite the Kremlin's desire to cut off Ukraine transit at the end of the 10-year contract that expired in December 2019.[29] Each of these examples illustrates that energy remains important in Russian grievances.

The Russo-Ukraine War Begins

In addition to playing a role in the objectives of the conflict and grievances between the parties, energy was also a key tool of conflict. Energy and water infrastructure shared by Russia and Ukraine was physically targeted by actors loyal to Ukraine early in the war. A brief recap of the conflict will help illustrate this point.

In February 2014, following several months of protests in Kiev, the then president Viktor Yanukovych fled the country, and a pro-Western government formed in his wake. In March 2014, the Crimean Peninsula, with a majority ethnic Russian population, voted to secede from Ukraine and join the Russian Federation. The vote, which received criticism due to alleged Russian interference, was not recognized by the international community or by Ukraine. Russia, however, recognized the referendum and annexed the peninsula. Shortly after the annexation, local separatists and Russian "little green men" launched a secessionist war in the Donbas region, which borders Russia and has an ethnic Russian majority. As the conflict dragged on—killing more than 10,300 and injuring nearly 24,000 between 2014 and October 2021—Ukrainian forces limited the territory held by the separatists but could not displace them entirely.[30] Figure 5-1 depicts the conflict areas in Crimea and the Donbas region in 2014.[31]

On November 20, 2015, with a hot war in the Donbas still underway, saboteurs loyal to Kiev blew up key pylons connecting Crimea to Ukraine's grid, destroying Crimea's access to Ukrainian electricity. Until this time, Crimea had continued to import 80 percent of its wintertime electricity from Ukraine, despite its decision to secede. The Ukrainian government did not claim credit for the sabotage, but also did not rush to fix the damage.

29. RadioFreeEurope/Radio Liberty, "Russia, Ukraine Reach Five-Year Gas-Transit Deal," RadioFreeEurope/Radio Liberty (website), December 31, 2019, https://www.rferl.org/a/long-russia-ukraine-reach- five-year-gas-transit-deal/30353000.html.

30. Council on Foreign Relations (CFR) Global Conflict Tracker, "Conflict in Ukraine," CFR (website), October 11, 2021, https://www.cfr.org/global-conflict-tracker/conflict/conflict-ukraine.

31. "Donbass Map - 2014," Global Security (website), October 11, 2021, https://www.globalsecurity.org/jhtml/jframe.html#https://www.globalsecurity.org/military/world/para/images/map-donbass-2014.jpg%7C%7C%7CDonbass%20Map%20-%202014.

Crimea's 1.9 million residents were plunged into darkness.[32] With no land connection to Crimea, Russia struggled to supply generators and establish replacement power for its new client republic. Crimea was still in crisis and largely in the dark when the first cyber-induced power outage hit Ukraine on December 23, 2015.

Figure 5-1. Russia-Ukraine conflict areas (2014)
(Map by Global Security)

The Cyber War Begins (BlackEnergy and KillDisk)

In the afternoon of December 23, 2015, electricity systems in three Ukrainian regions, including Kiev, began shutting down. Despite desperate efforts by operators, malicious actors were able to take control remotely of more than 50 substations. Many of the control center operators watched helplessly as the malware invalidated their passwords and as remotely controlled

32. Anna Shamanska, "Why Ukraine Supplies Electricity to Crimea, and Why It Stopped," RadioFreeEurope/RadioLiberty (website), November 24, 2015, https://www.rferl.org/a/ukraine-crimea-power-supply-electricity-explainer/27384812.html.

cursors moved across their screens.[33] By the end of the attack, 130 megawatts of generating capacity dropped from the grid, leaving some 225,000 customers across Ukraine without electricity.[34] The situation was made worse by a denial of telephone service attack that used automated systems to overload the phones, making it impossible for customers to report outages, and making it difficult for power stations to communicate with each other. By pulling all supervisor control and data acquisition (SCADA) systems offline and functioning in manual mode, operators restored power in one to six hours, but the substations were not fully operational for months.[35]

The attackers managed to take over systems by way of remote logons, obtained through spear phishing and breaching of data systems.[36] Expert analysis noted that the attack began six months prior with relatively sophisticated e-mail spear-phishing attacks that appeared to be official correspondence from the Ukraine Energy Ministry. At least one user at each of the six targeted distribution centers was taken in by the spear phishing, which then released a BlackEnergy macro into key systems when the file was opened.[37] Once within the systems, these macros completed network reconnaissance, siphoned off credentials and passwords, and accessed the control center, providing the attackers full information about the system. At the appointed time, the system attackers took control, opening breakers to shut the systems down. The entire operation took approximately 60 minutes. After bringing down the grid, attackers used KillDisk malware to destroy some systems in each of the three master stations.[38] The attack focused on the distribution sector, which is the most decentralized sector of the grid. More than two months later, critical devices at 16 substations remained damaged, and breakers had to be controlled manually.[39]

In previous months, BlackEnergy and KillDisk malware had been found in and had done damage to Ukrainian government agencies, television

33. Kim Zetter, "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid," Wired (website), March 3, 2016, https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/.

34. David E. Whitehead et al., "Ukraine Cyber-induced Power Outage: Analysis and Practical Mitigation Strategies," in 70th Annual Conference for Protective Relay Engineers (Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2017), 1–2, https://doi.org/10.1109/CPRE.2017.8090056.

35. Whitehead et al., "Ukraine Cyber-induced Power Outage," 3.

36. Joe Slowick, *CRASHOVERRIDE: Reassessing the 2016 Ukraine Electric Power Event as a Protection-Focused Attack* (Hanover, MD: Dragos, August 15, 2019), 2–3, https://www.dragos.com/wp-content/uploads/CRASHOVERRIDE.pdf.

37. Whitehead et al., "Ukraine Cyber-induced Power Outage," 2.

38. Whitehead et al., "Ukraine Cyber-induced Power Outage," 3.

39. Zetter, "Inside the Cunning."

networks, and the state rail company, and had also been found in Polish energy firms and even in some power and water utilities in the United States.[40] The use of this malware in the electricity grid was still shocking because it represented sophisticated long-term planning and is widely understood to be the first instance of a hack successfully taking down a power grid. Although some argued that the cyberattack was retribution for Ukraine's role in denying electricity to Crimea, systems intrusions had already begun in March, prior to the Crimean blackout. Evidence suggests, however, that the preexisting attack plan may have been rushed to implementation after the physical attack on the Crimean grid.[41]

The Cyber War Escalates (CrashOverride)

The second attack against Ukraine's electricity grid occurred on December 17, 2016, when, just before midnight, a cyberattack on the electric transmission station outside of Kiev took down 20 percent of its capacity. This attack targeted a transmission facility rather than distribution facilities as the previous attack had done.[42] The automated attack opened every circuit breaker in a key transmission station.[43] Ukraine's electricity company, *Ukenergo,* and Ukrainian cybersecurity experts asserted the attack was similar to the first one, but subsequent analysis revealed something more concerning.

The attack had been engineered to inflict physical damage on the electrical equipment. A Dragos cybersecurity team researching the incident concluded that the power outage was supposed to be the first stage of a larger, more ambitious attack. While the power was down, the attackers had launched an attack on the protective relays, apparently to disable fail-safe devices and make it dangerous to restart the grid. The Dragos report concludes the hackers had intended to cause extensive damage *after* the power was restored.[44] Based on the previous response in 2015, the attackers expected that company engineers would rush to restart the systems manually. By disabling the fail-safe devices, the attackers intended for workers who were restoring the system to trigger a current overload, which would have caused potentially

40. Andy Greenberg, "How an Entire Nation Became Russia's Lab for Cyberwar," Wired (website), June 28, 2017, https://www.wired.com/story/russian-hackers-attack-ukraine/.

41. Zetter, "Inside the Cunning."

42. Kim Zetter, "The Ukrainian Power Grid Was Hacked Again," Motherboard Tech by Vice (website), January 10, 2017, https://www.vice.com/en/article/bmvkn4/ukrainian-power-station-hacking-december-2016-report.

43. Andy Greenberg, "New Clues Show How Russia's Grid Hackers Aimed for Physical Destruction," Wired (website), September 12, 2019, https://www.wired.com/story/russia-ukraine-cyberattack-power-grid-blackout-destruction/.

44. Slowik, *CRASHOVERRIDE*, 9.

catastrophic damage resulting in long disruptions to energy transmission, damaging power generation assets, and injuring workers.[45] This type of cyber disruption is relatively rare, and, as this instance illustrates, it is difficult to implement successfully but threatens severe consequences. It is notable that this early use of a cyber weapon was both state-sponsored and aimed at energy infrastructure, specifically targeting the transmission and generation sectors, on which the whole grid depends. See chapter 14, which discusses this attack among 10 major cyber incidents.

Although the attack failed, largely due to attackers' errors, the level of ambition and novel approach is notable. It made use of software to manipulate industrial control systems, rather than using the manual interaction of hackers in real time with the systems, as had been done in 2015.[46] The attackers used a malware known as CrashOverride. This malware targeted protective relays, which automatically open circuit breakers if dangerous conditions threaten to damage transformers. The attackers were aware of a security flaw in the relays that made it possible to put them into firmware update mode, which would render the relays unusable until rebooted manually.[47] The malware failed only because the attackers failed to enter the Internet protocol addresses of the protective relays properly.[48]

Even if these addresses had been input correctly, it is not certain that the attack would have succeeded.[49] Malware-induced physical sabotage has been attempted "in the wild" only two other times: the Stuxnet attack on Iranian nuclear facilities in 2009 and the Triton/Trisis attack on the Saudi Petro Rabigh oil refinery in 2017. The Triton/Trisis attack, linked to Moscow's Central Scientific Research Institute of Chemistry and Mechanics, shut down the Saudi plant but did not damage it.[50] In other words, the CrashOverride attackers used cutting edge, unproven methods, though they did have some evidence that their approach might work. In 2007, tests at the Idaho National Laboratory demonstrated that corrupting the protective relay in an electricity transmission system could destroy a 27-ton generator.

45. Greenberg, "New Clues."
46. Slowik, *CRASHOVERRIDE*, 3.
47. Greenberg, "New Clues."
48. Slowik, *CRASHOVERRIDE*, 12.
49. Slowik, *CRASHOVERRIDE*, 12.
50. Greenberg, "New Clues."

Mike Assante, the project lead on this secret project, served on the team investigating the Ukraine attack and noticed the similarities in approach.[51]

It is not surprising that the seriousness of the attack was not initially recognized in Ukraine since it came amid a wave of hundreds of other cyberattacks. In July 2016, a massive phishing campaign targeted government organizations. Then in December 2016, attacks unfolded against the Ukrainian Ministry of Finance, the State Treasury, the Pension Fund, the State Administration of Railway Transport, and the Defense Ministry.[52] Afterwards, the then President Petro Poroshenko reported that in the last two months of 2016, hackers targeted government institutions some 6,500 times, with the recent attacks on the State Treasury preventing the timely payment of salaries and pensions.[53] Denial-of-service attacks hit the Defense Ministry a week after the State Treasury attacks and took down its website in an apparent effort to disrupt updates about the conflict in Donbas.[54]

Collateral Damage: A Cyberattack on the Ukrainian Economy (NotPetya)

Unlike the first two cyberattacks, the third attack did not specifically target energy infrastructure, though energy infrastructure became involved as the crisis unfolded. In February 2017, the US National Security Agency (NSA) warned Microsoft that the NSA's EternalBlue code—which exploited flaws in a Windows protocol, making it possible to take over a vulnerable computer remotely—had been stolen. Microsoft issued a patch in March, but there were still computers all over the world that had not yet been updated.[55] WannaCry in April and NotPetya in rapid succession made efficient use of EternalBlue in their malware.

NotPetya was unleashed in Ukraine in June 2017, just before the Constitution Day holiday. The attackers deliberately infected M.E. Doc, a business software widely used in Ukraine. When customers went to the website for an update, the malware accessed personal computers all

51. Andy Greenberg, "How 30 Lines of Code Blew Up a 27-ton Generator," Wired (website), October 23, 2020, https://www.wired.com/story/how-30-lines-of-code-blew-up-27-ton-generator/.

52. Zetter, "Ukrainian Power Grid."

53. Natalia Zinets, "Ukraine Hit by 6,500 Hack Attacks, Sees Russian 'Cyberwar'" BBC (website), December 29, 2016, https://www.reuters.com/article/us-ukraine-crisis-cyber-idUSKBN14I1QC.

54. Reuters, "Ukraine's Defence Ministry Says Website Hit by Cyber Attack," Reuters (website), December 13, 2016, https://www.reuters.com/article/us-ukraine-crisis-cyber/ukraines-defence-ministry-says-website-hit-by-cyber-attack-idUSKBN1421YT.

55. Andy Greenberg, "Hold North Korea Accountable for WannaCry—and the NSA, Too," Wired (website), December 19, 2017, https://www.wired.com/story/korea-accountable-wannacry-nsa-eternal-blue/.

over Ukraine and beyond. NotPetya became the fastest-propagating malware ever seen.[56] Although energy systems were not the primary target of this attack, Ukraine's transmission system operator, *Ukrenergo*—still recovering from the cyberattacks of 2015 and 2016—was hit just before it completed safety upgrades. *Kyivenergo*, Ukraine's heat and energy company, was also hit along with five other power companies.[57]

Even though NotPetya demanded a payment of $300 in bitcoin from owners of computers it infected, it was not ransomware.[58] Instead, the malware permanently destroyed data even as its victims were reading the ransom note, erasing master boot systems and causing computers to be unable to find their operating systems.[59] Ukraine was hit particularly hard. In addition to the energy infrastructure, multiple national media outlets, four hospitals, two airports, and more than 22 banks, automatic-teller machines, and card payment systems were damaged in the attack. An estimated 10 percent of all the computers in the country were destroyed by the NotPetya attack.[60]

The damage, however, was not limited to Ukraine, but impacted organizations in an estimated 65 countries as well.[61] A White House assessment estimates more than $10 billion in damages.[62] Perhaps the most famous victim of NotPetya was Maersk, a major global maritime shipping company. Maersk became infected through a single laptop belonging to a finance executive in Odessa. Maersk reported that NotPetya infected most of its network within seven minutes and destroyed 49,000 laptops in the organization while paralyzing global shipping.[63] Among the companies that reported

56. Andy Greenberg, "The Untold Story of NotPetya, the 'Most Costly Cyberattack in History,'" Wired (website), August 22, 2018, https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/.

57. Greenberg, "Untold Story of NotPetya."

58. Riley Griffin, Katherine Chiglinsky, and David Voreacos, "Was It an Act of War? That's Merck Cyber Attack's $1.3 Billion Insurance Question," Bloomberg (website), December 3, 2019, https://www.bloomberg.com/news/features/2019-12-03/merck-cyberattack-s-1-3-billion-question-was-it-an-act-of-war.

59. Greenberg, "Untold Story of NotPetya."

60. Greenberg, "Untold Story of NotPetya."

61. Jai Vijayan, "3 Years after NotPetya, Many Organizations Still in Danger of Similar Attacks," *Dark Reading Newsletter* (website), June 30, 2020, https://www.darkreading.com/threat-intelligence/3-years-after-notpetya-many-organizations-still-in-danger-of-similar-attacks/d/d-id/1338200.

62. Greenberg, "Untold Story of NotPetya."

63. Eric Parizo, "Maersk CISO Says NotPetya Devastated Several Unnamed US firms," *Dark Reading Newsletter* (website), December 9, 2019, https://www.darkreading.com/omdia/maersk-ciso-says-notpeyta-devastated-several-unnamed-us-firms.

losses, the damages are staggering: Merck lost $870 million and suffered 30,000 computers and 7,500 servers crippled by the attack; FedEx/TNT lost $400 million; Saint-Gobain lost $380 million; and Maersk lost $300 million.[64] See chapter 4 for an analysis on the potential future risks to the US Transportation Command—and thus, to US and NATO force project and military mobility—based on NotPetya's impact on global shipping.

While cyber experts are divided on whether NotPetya's impact exceeded the intentions of its creators, they generally describe it as the closest example of cyber war to date. As cybersecurity expert Kenneth Geers notes, "This was the most damaging attack in history, of a scale and a cost that would far exceed a missile fired from the Donbas into Kiev."[65] The rhetoric regarding cyber war is more than rhetoric: Merck's $1.3 billion insurance claim was declined on the grounds that the attack was an act of war, for which the company was not covered.[66] Some experts, including Ukraine's cybersecurity service, argue that the attack also served as a cleanup effort, enabling hackers to destroy evidence of espionage or reconnaissance.[67] Other experts, such as Cisco's Craig Williams, conclude that the message—that it is dangerous to do business in Ukraine—was powerfully and deliberately delivered.[68] Regardless of whether the attackers had anticipated the impact of NotPetya, it did shift the balance as nations began to attribute the cyberattacks to Russia.

Attribution Evolves

Attribution—the act of identifying who is responsible for a cyberattack—is notoriously difficult, making it an attractive option in hybrid conflict. Most cyber actors cover their tracks, routing their attacks through other countries or using techniques known to be associated with a different criminal group to mislead investigators. Within two weeks of the first attack, the United States sent an expert team to Ukraine to better understand this type of attack and develop strategies for protecting against it.[69] The team concluded that the attack showed evidence of some state sponsorship and

64. Griffin, Chiglinsky, and Voreacos, "Was It an Act of War?," (2019); and Greenberg, "Untold Story of NotPetya."

65. Laurens Cerulus, "How Ukraine Became a Test Bed for Cyberweaponry," Politico (website), February 14, 2019, https://www.politico.eu/article/ukraine-cyber-war-frontline-russia-malware-attacks/.

66. Griffin, Chiglinsky, and Voreacos, "Was It an Act of War?," (2019).

67. Greenberg, "Untold Story of NotPetya."

68. Greenberg, "Untold Story of NotPetya."

69. "ICS Alert: Cyber-Attack against Ukrainian Critical Infrastructure," Cybersecurity and Infrastructure Security Agency, n.d., last revised July 20, 2021, https://us-cert.cisa.gov/ics/alerts/IR-ALERT-H-16-056-01.

support. The United States, however, was much less eager to attribute an attack to Russia than it had been to attribute previous attacks to North Korea, despite a persistent rise in such attacks. See chapter 4 for fascinating detail on evidence of other Russian-led cyber intrusions and attacks.

Ukraine had long argued that each of the three cyberattacks described above bore similar "handwriting" and techniques.[70] Ukrainian experts identified Sandworm early on as the team responsible for these and many other cyberattacks on Ukrainian soil. As international experts gathered more information, they, too, began to believe in the likelihood of state sponsorship.

Within six months of the attack, the US Central Intelligence Agency allegedly attributed the NotPetya attacks to Russia with high confidence, but this report was classified.[71] Not until February 2018 did US intelligence agencies confirm the Russian military's role in launching NotPetya. The US statement came a few hours behind British and Danish government statements, both of which had already condemned the Russian military's role in the attack.[72] The statement from the then White House press secretary included the following accusation:

> It was part of the Kremlin's ongoing effort to destabilize Ukraine and demonstrates ever more clearly Russia's involvement in the ongoing conflict. This was also a reckless and indiscriminate cyber-attack that will be met with international consequences.[73]

Although clear international consequences were arguably not imposed, evidence continued to mount as US investigators did their work. By October 2020, the US Department of Justice unsealed an indictment against the Sandworm hacking group, accusing each of the six members by name and clearly identifying them as Russian intelligence officers. According to the statement, "no country has weaponized its cyber capabilities as maliciously or irresponsibly as Russia, wantonly causing unprecedented damage to pursue

70. Andy Greenberg, "Petya Ransomware Epidemic May Be Spillover from Cyberwar," Wired (website), June 28, 2017, https://www.wired.com/story/petya-ransomware-ukraine/.

71. Ellen Nakashima, "Russian Military Was behind 'NotPetya' Cyberattack in Ukraine, CIA Concludes," *Washington Post*, January 12, 2018, https://www.washingtonpost.com/world/national-security/russian-military-was-behind-notpetya-cyberattack-in-ukraine-cia-concludes/2018/01/12/048d8506-f7ca-11e7-b34a-b85626af34ef_story.html.

72. Greenberg, "Untold Story of NotPetya."

73. Sarah Huckabee Sanders, "Statement from the White House Press Secretary," US Embassy and Consulates in Russia (website), February 15, 2018, https://ru.usembassy.gov/statement-white-house-press-secretary-021518.

small tactical advantage and to satisfy fits of spite."[74] Attacks against the Ukrainian government and its critical infrastructure were the first offense described in the statement, followed by explicit reference to BlackEnergy, Industroyer, KillDisk, and NotPetya.[75]

There are multiple reasons why the United States was slow to blame Russia for cyberattacks. Some note that public blame must logically be followed with retaliation—and there is little consensus on appropriate retaliation—while others focus on the fact that the United States is also very involved in covert cyber operations.[76] The origins of EternalBlue highlight this problem as NSA alerted Microsoft to a vulnerability in its system after the leak of EternalBlue, which had been developed to exploit that vulnerability.[77]

Learning from Ukraine: Improving Infrastructure Safeguards

When Ukraine joined the EU Energy Community in 2011, it committed to reforming its energy sector to comply with European standards and adopting EU internal energy market legislation.[78] In exchange, Ukraine has received extensive support for energy sector reform, and is increasingly integrated into European energy systems. In electricity, joining European energy markets poses a particular problem: Ukraine's electricity grid remains synchronized with the Russian grid. This is not unique to Ukraine; the Baltic states also remain synchronized with the Russian and Belarusian grids. The effort to move the Baltic states fully into the European grid began in 2009 and has a target date for completion of 2025.[79] This process remains complex and delicate, as it requires transformation of infrastructure, fundamental shifts in electricity trade, and acceptance of higher prices.[80]

74. Department of Justice, "Six Russian GRU Officers Charged.".

75. Department of Justice, "Six Russian GRU Officers Charged."

76. Dustin Volz and Sarah Young, "White House Blames Russia for 'Reckless' NotPetya Cyber Attack," Reuters (website), February 15, 2018, https://www.reuters.com/article/us-britain-russia-cyber-usa/white-house-blames-russia-for-reckless-notpetya-cyber-attack-idUSKCN1FZ2UJ.

77. Greenberg, "Hold North Korea Accountable."

78. "Who We Are," EU Energy Community (website), n.d., accessed September 28, 2021, https://www.energy-community.org/aboutus/whoweare.html.

79. European Commission, "Baltic Energy Market Interconnection Plan," European Commission Energy, last updated October 29, 2021, https://ec.europa.eu/energy/topics/infrastructure/high-level-groups/baltic-energy-market-interconnection-plan_en.

80. Richard Morningstar, "The Geopolitics of Electricity Security in Northeast Europe," Atlantic Council, streamed live on June 28, 2021, You Tube video, 1:09:53, https://www.youtube.com/watch?v=uL5ppzQzVk0.

Rather than discouraging the EU, the cyberattacks against Ukraine proved to be a catalyst for a concerted effort to move Ukraine further into the EU energy space. In June 2017, the EU executed an agreement with Ukraine and Moldova to work toward full synchronization of the grids following each country's removal of itself from the Russian grid. The set date for completion is 2023. The goals of integrating Ukraine into the European grid are to enhance reliability and security of supply, increase competitiveness, and improve grid resilience as well as to diversify the energy mix for both Ukraine and Europe. The cost of the synchronization is estimated at $400 million, while the benefits are estimated to exceed $1 billion annually.[81]

As the first major step toward Ukrainian-EU grid integration, one 2,300-megawatt coal power plant at Burshtyn Island was disconnected from the Ukrainian grid and joined to the European grid.[82] The success of this project led to proposals for an energy bridge which would allow one nuclear power plant, Khmelnitski-2, to be moved into the European grid as well.[83] After initial delays, the tender for this project was awarded in August 2019 to the Ukraine Power Bridge Company, a consortium including Westinghouse, EDF, Polenergia International, and the Hungarian national energy company MVM.[84]

By bringing Ukraine into the European grid, the EU will obtain more insight into the Ukrainian systems, ability to train Ukrainian operators, and access to additional nuclear power at a time when some EU states—most notably Germany—are moving away from nuclear power. Ukraine, which ranks seventh among world nuclear energy producers, remains strongly pronuclear, with 15 commercially operating nuclear power units.[85] The European Bank for Reconstruction and Development (EBRD)

81. Sandeep Kohli, "Ukraine: Facilitating Power System Integration with Europe Project," World Bank (website), April 3, 2020, https://documents.worldbank.org/en/publication/documents-reports/documentdetail/989951585894725327/concept-project-information-document-pid-ukraine-facilitating-power-system-integration-with-europe-project-p171980.

82. International Energy Agency (IEA), *Ukraine Energy Profile* (Paris: IEA, April 2020), 33, https://www.iea.org/reports/ukraine-energy-profile.

83. "Nuclear Power in Ukraine," World Nuclear Association (website), n.d., last updated September 2021, https://www.world-nuclear.org/information-library/country-profiles/countries-t-z/ukraine.aspx.

84. "Energy Bridge Has Geopolitical Significance, Says Polenergia," World Nuclear News (website), December 4, 2020, https://www.world-nuclear-news.org/Articles/Energy-Bridge-has-geopolitical-significance,-says.

85. IEA, *Key World Energy Statistics 2020* (Paris: IEA, August 2020), 19, https://www.iea.org/reports/key-world-energy-statistics-2020.

has provided significant assistance for reactor security upgrades.[86] Supporters of EBRD's assistance argue it is the EU's best chance to improve safety and security in Ukrainian energy while simultaneously making nuclear power available to a decarbonizing Europe.

The grid integration is very ambitious both economically and physically, and critics note that Ukraine would have to delink from Russia before it can join the European grid.[87] Economically, such a shift would cut Ukraine off from the option of importing electricity from Belarus, which is often cheaper.[88] From an infrastructure perspective, Ukraine's transmission network is massive and inefficient. EBRD invested approximately $124 million in improving the grid beginning in 2015, but it is estimated that $5.1 billion additional investment will be required.[89] From an energy security perspective, this move greatly enhances Ukraine's energy security, but due to the highly networked nature of electricity, it will pose new threats to the European grid.

The problem of Ukraine's cyber vulnerability is a matter of concern. The Institute of Electrical and Electronics Engineers published a series of recommendations based on Ukraine's experience, including a framework for better protecting substations against attacks like CrashOverride. EU supporters of electricity integration highlight the extent to which best practices can be shared with Ukraine, both in nuclear power safety and in cybersecurity. They also note that since the 2014 Association Agreement with the EU, Ukraine has become so economically entangled with Europe that the risk of a "Ukraine contagion" is already apparent. Data flows and interactions with Ukrainian Internet networks already threaten Europe, as evidenced by the NotPetya attack.[90] There is little question, however, that integration of Ukrainian electricity into the EU system represents exposure to new cyberattack possibilities for Europe.[91]

86. "Ukraine Prepares to Reduce Output during Pandemic," World Nuclear News (website), April 30, 2020, https://world-nuclear-news.org/Articles/Ukraine-prepares-to-reduce-output-during-pandemic.

87. Olena Holubeva, "Ukraine's Disconnection from Power Grids of Russia and Belarus: Prices and Consequences," 112, UA International (website), April 28, 2021, https://112.international/politics/disconnection-of-ukraine-from-power-grids-of-the-russia-and-belarus-prices-and-consequences-60945.html.

88. Holubeva, "Ukraine's Disconnection."

89. Anton Antonenko et al., "Reforming Ukraine's Energy Sector: Critical Unfinished Business," Carnegie Europe (website), February 6, 2018, https://carnegieeurope.eu/2018/02/06/reforming-ukraine-s-energy-sector-critical-unfinished-business-pub-75449.

90. Cerulus, "How Ukraine Became a Test Bed."

91. Heng Chuan Tan et al., "Tabulating Cybersecurity Solutions for Substations: Towards Pragmatic Design and Planning," 2019 IEEE PES Innovative Smart Grid Technologies – Asia (Piscataway, NJ: IEEE, 2019), 1020–1022, https://doi.org/10.1109/ISGT-Asia.2019.8881706.

Improving Ukraine: Vulnerabilities

Cyber Vulnerabilities

The institute's study of Ukraine, described previously, identifies enforcement of polices and protocols as an important measure that is difficult to subvert but which has the downside of being intrusive and carrying high deployment, operating, and maintenance costs.[92] Enforcement remains a key challenge in the Ukraine context, where even the most minimal policies and protocols struggle to gain traction. A majority of Ukraine's computers run on pirated software that does not receive standard security patches and that may be corrupted.[93] As late as 2016, an estimated 82 percent of all software used in Ukraine was unlicensed.[94]

The problem extends to official use of software as well as private and commercial use. In 2020, Ukraine was on the 10-country "Priority Watch List" of the Office of the US Trade Representative's Special 301 Report, which identifies countries with problematic intellectual property (IP) rights behavior. In this report, "widespread use of unlicensed software by Ukrainian government agencies" was enumerated among the top three concerns.[95] Similarly, a 2020 biannual EU report on IP rights in Ukraine notes that little progress has been made in the area of IP protection and enforcement, causing the EU ongoing concern.[96] Although the government of Ukraine is improving the laws governing IP, moving away from illegal software will take time for the society, including the government, to achieve. Inattention to safe software is one of many problems in Ukrainian tech. In 2018, the British tech research firm Comparitech ranked Ukraine the 10th-least cyber secure country out of the 60 nations researched, estimating 28 percent of Ukrainian computers and 11 percent of phones were infected with malware.[97] Given this context, Ukraine will be challenged to implement important

92. Tan et al., "Tabulating Cybersecurity Solutions," 1019.

93. Cerulus, "How Ukraine Became a Test Bed."

94. "Software Piracy: Why You Shouldn't Get Scared of Outsourcing to Ukraine," IT Outsourcing Review (website), February 7, 2017, https://outsourcingreview.org/software-piracy-why-you-shouldnt-get-scared-of-outsourcing-to-ukraine/.

95. Office of the United States Trade Representative, *2020 Special 301 Report* (Washington, DC: Office of the US Trade Representative, April 2020), 58, https://ustr.gov/sites/default/files/2020_Special_301_Report.pdf.

96. "EU Report on Intellectual Property Rights Highlights Developments in Ukraine," EU Neighbors East (website), January 13, 2020, https://www.euneighbours.eu/en/east/stay-informed/news/eu-report-intellectual-property-rights-highlights-developments-ukraine.

97. "Ukraine Cybersecurity Assistance," International Trade Administration (website), September 30, 2020, https://www.trade.gov/market-intelligence/ukraine-cybersecurity-assistance.

cybersecurity measures without significant shifts in its security culture. See chapter 14 for an overview of recommended cybersecurity measures.

Recognizing the cyber threats to energy infrastructure, *Ukrenergo* established a cybersecurity operations center in 2019, following the 2018 effort of the National Security and Defense Council of Ukraine to coordinate all of the country's cybersecurity initiatives.[98] In addition, *Ukrenergo* announced in early 2018 that it was planning a $20 million cyber defense system for its operations, with a completion target of 2020.[99] Meanwhile, Russian-based attacks on energy infrastructure continue, though none have been as spectacularly successful as the efforts in 2015–18. Of greatest concern is the cybersecurity of Ukraine's nuclear power. In 2018, a group known collectively as the Cyber Alliance—comprised of four "hacktivist" groups called CyberHunta, Falcons Flame, Trinity and RUH8—controversially tested their own country's systems in an effort to find vulnerabilities. In their reports to the press, Cyber Alliance claimed to have successfully probed Energoatom and found vulnerabilities that would easily allow hackers to enter the system of one of its nuclear facilities. See chapter 3 for several striking examples of cyber challenges to nuclear power plants and the cascading effects such attacks can cause.[100]

Nuclear Vulnerabilities

According to its official documents, Ukraine intends to remain approximately 50 percent reliant on nuclear energy even as demand for electricity rises. Ukraine signed nuclear energy cooperation agreements with the EU and hopes to export nuclear-supplied electricity to the EU in growing quantities. The updated Ukraine Energy Strategy to 2030 plans for 5,000–7,000 megawatts of new nuclear power at an estimated cost of $25 billion.[101] The nuclear energy sector raises particular concerns in terms of CI and cybersecurity.

98. "Ukrenergo Develops a Roadmap for Cyber Defense Development and Sets Up a Security Operations Centre," Ukrenergo (website), March 19, 2019, https://ua.energy/main-events/ukrenergo-develops-a-roadmap-for-cyber-defense-development-and-sets-up-a-security-operations-centre/; and Christopher Miller, "What's Ukraine Doing to Combat Russian Cyberwarfare? 'Not Enough,'" RadioFreeEurope/Radio Liberty (website), March 7, 2018, https://www.rferl.org/a/ukraine-struggles-cyberdefense-russia-expands-testing-ground/29085277.html.

99. "Ukraine Power Distributor Plans Cyber Defense System for $20 Million," Reuters (website), February 6, 2018, https://www.reuters.com/article/us-ukraine-cyber-ukrenergo/ukraine-power-distributor-plans-cyber-defense-system-for-20-million-idUSKBN1FQ1TD.

100. Miller, "What's Ukraine Doing."

101. "Nuclear Power in Ukraine."

The Nuclear Threat Initiative Nuclear Security Index, which bases its analysis on data from the International Atomic Energy Agency (IAEA), provides cybersecurity scores for nuclear power facilities in 47 countries. In its 2016 ratings, Ukraine received only one point (of four possible points), placing it in the bottom half of countries.[102] The index also measures other aspects of nuclear power plant security—on which Ukraine scores significantly better—but its risk environment is identified as unfavorable due to pervasive corruption, low political stability, and ineffective governance.[103]

There is also some concern regarding fuel for the reactors. Nuclear fuel supply is less intensively networked than natural gas. Plants can store a year's worth of fuel on-site, and most have dry cask storage for nuclear waste. While Ukraine still receives most of its nuclear services and nuclear fuel from Russia, it is reducing this dependence by buying American nuclear fuel. In 2016, Ukraine procured 40 percent of its fuel from Westinghouse, and, with US encouragement, expects to expand these purchases over time.[104] Future EU priorities regarding nuclear power are unclear, as European states are divided over its desirability, but Ukraine is determined to play a role in the European grid with its nuclear supplied electricity. The EU, for its part, has already been very involved in Ukraine's nuclear energy future, as the next section will demonstrate.

Improving Ukraine: Assistance

Although the United States was slow to attribute the attacks to Russia, Ukraine quickly became the place to study cyberattack. Ukrainian authorities have been consistently willing to provide cyber intelligence in exchange for assistance in fending off further cyberattacks.[105] Ukraine's information and intelligence sharing with the EU, NATO, and the United States has been well-developed. Even before the cyberattacks on energy infrastructure, the United States and NATO engaged in Ukrainian cybersecurity. NATO established its Trust Fund on Cyber Defence for Ukraine in 2014, led by Romania. The project, which was designed to last for two

102. Alexandra Van Dine, Michael Assante, and Page Stoutland, *Outpacing Cyber Threats: Priorities for Cybersecurity at Nuclear Facilities* (Washington, DC: Nuclear Threat Initiative, 2016), 14, https://media.nti.org/documents/NTI_CyberThreats__FINAL.pdf.

103. "NTI Index for Ukraine," Nuclear Threat Initiative (website), n.d., accessed June 21, 2021, https://www.ntiindex.org/country/ukraine/.

104. "Nuclear Power in Ukraine," World Nuclear Association, (2021).

105. Cerulus, "How Ukraine Became a Test Bed."

years, included the establishment of laboratories that could investigate cybersecurity incidents.[106]

Following the cyberattacks on infrastructure and the US-led investigations, the United States began investing directly in Ukraine's cybersecurity, convening regular US-Ukraine Cybersecurity Dialogues beginning in September 2017. The US Congress adopted the "Ukraine Cybersecurity Cooperation Act of 2017" and set a course for ongoing support and consultation.[107] The Cooperation Act required the United States to help with advanced security protection on government computers, protection of critical infrastructure, and building capacity. The Senate followed with similar legislation, the "Ukraine Cybersecurity Cooperation Act of 2018."[108]

From 2017–20, the United States convened three Cybersecurity Dialogues with Ukraine.[109] The United States announced financial support at the dialogues, with the Department of State pledging $10 million in 2017 and an additional $8 million in 2020 in support of Ukraine's cybersecurity capabilities. The 2020 funds are part of a US Agency for International Development cybersecurity project which is committed to investing up to $38 million over four years to support legal and regulatory reform, development of the cyber workforce, and private sector engagement.[110] The assistance supports Ukraine's efforts to adhere to recently signed treaties to improve the security of its cyberspace.

NATO's efforts to combat hybrid threats rose largely in response to Russian aggression along NATO's eastern flank. See chapter 4 for a more comprehensive explanation of NATO's efforts to counter hybrid threats. Since the 2014 Russia-Ukraine conflict began, NATO has intensified its cooperation with Ukraine. In 2020, NATO designated Ukraine

106. NATO Trust Fund, "Ukraine: Cyber Defense," NATO Trust Fund (website), June 2016, https://www.nato.int/nato_static_fl2014/assets/pdf/pdf_2016_07/20160712_1606-trust-fund-ukr-cyberdef.pdf.

107. Ukraine Cybersecurity Cooperation Act of 2017, H.R. 1997 (2017), https://www.congress.gov/bill/115th-congress/house-bill/1997.

108. Ukraine Cybersecurity Cooperation Act of 2018, S.2455 (2018), https://www.congress.gov/bill/115th-congress/senate-bill/2455.

109. "Embassy Statement on the First US-Ukraine Bilateral Cyber Dialogue," US Embassy in Ukraine (website), September 29, 2017, https://ua.usembassy.gov/embassy-statement-first-us-ukraine-bilateral-cyber-dialogue/; and Office of the Spokesperson, "Second U.S.-Ukraine Cybersecurity Dialogue," Department of State (website), November 5, 2018, https://ua.usembassy.gov/second-u-s-ukraine-cybersecurity-dialogue/.

110. Office of the Spokesperson, "The United States and Ukraine Hold Third Cyber Dialogue," Department of State (website), March 3, 2020, https://ua.usembassy.gov/the-united-states-and-ukraine-hold-third-cyber-dialogue/.

as an Enhanced Opportunities Partner, and Ukraine, in turn, references the aims of EU and NATO membership in its 2020 national security strategy.[111]

Not surprisingly, however, the greatest assistance to Ukraine has come from the EU. In 2019, EU institutions constituted the lead source of all overseas development assistance to Ukraine, providing $413 million. In addition, Germany and Poland made bilateral contributions of $205 million and $81 million respectively, compared to $198 million by the United States.[112] Europe's leadership of assistance to Ukraine is reflected in its support of critical infrastructure as well. Ukraine joined the EU Energy Community in 2011, committing to reform its energy sector to comply with European standards and adopting EU internal energy market legislation.[113] In exchange, the EU provides extensive support to Ukraine, primarily through workshops and projects under the European Programme for Critical Infrastructure Protection (EPCIP). EPCIP aims to harmonize approaches to CISR in the transport and energy sectors, with an emphasis on infrastructure with a transboundary effect. See chapter 10 for in-depth analysis of EPCIP and the broader EU policy framework for CISR.

The EU has already demonstrated a long-term commitment to Ukrainian nuclear energy and completed several extremely ambitious investments in Ukraine and energy security. Most notably, the Chernobyl Shelter Implementation Plan (2010–19), spent €2.1 billion to contain the radioactive remains of the destroyed Chernobyl Reactor 4 fully.[114] The EU also financed the completion of three nuclear reactors in Ukraine—Khmelnitski-2, Rovno-4, and Zaporizhzhya-6—that had been begun in the Soviet era.[115]

The EU's decision to include Ukraine in the European grid, which is facilitated by the European Network of Transmission System Operators for Electricity (ENTSO-E), is ambitious but has already borne some fruit. In February 2021, Ukraine's upper legislative body supported a draft law making it possible for *Ukrenergo* to be certified under EU rules—

111. Natalya Belitser, "Some Thoughts on the Adoption of Ukraine's National Security Strategy," US-Ukraine Foundation, September 24, 2020, https://usukraine.org/news/articles/some-thoughts-on-the-adoption-of-ukraines-national-security-strategy/NTk2MTc=/; and "Relations with Ukraine," NATO (website), n.d., last updated August 27, 2021, https://www.nato.int/cps/en/natolive/topics_37750.htm.

112. US Agency for International Development (USAID), "Development Cooperation Landscape – Ukraine," USAID Foreign Aid Explorer (website), n.d., accessed June 22, 2021, https://explorer.usaid.gov/donor/ukraine.

113. "Who We Are," EU Energy Community.

114. Axel Reiserer, "Keys Handed Over for Chernobyl New Safe Confinement," European Bank for Reconstruction and Development (website), July 10, 2019, https://www.ebrd.com/news/2019/keys-handed-over-for-chernobyl-new-safe-confinement.html.

115. "Nuclear Power in Ukraine."

an essential component of the ongoing effort to move the entire Ukrainian grid into synchronization with the ENTSO-E grid.[116] *Ukrenergo* has also succeeded in transferring datasets to ENTSO-E and has completed the first stage of design of a new SCADA system in cooperation with the German conglomerate Siemens.[117] Improving the human capacity, particularly improving communications and building trust, is essential for Ukraine, and such joint projects are beginning to change the culture for the better. See chapter 3 for a useful discussion of the importance of human capacity building in countering cyber threats.

A Key Vulnerability Persists

Ukraine has been an important "sandbox" in which other nations have learned about cyberattacks. It has also been described as a "free-fire zone" in which Russia has been able to test new cyber weapons with little restraint. Often ignored by Western commentary is the extent to which Ukraine itself has been a leader in cyberattacks. Ukraine has an exceptionally high percentage of software developers, ranking first in the world when measuring software developers per 1,000 inhabitants.[118] Given this percentage and the profound asymmetry of the ongoing conflict with Russia, it should not be surprising that pro-Kiev forces engage in extensive cyberattacks against Russia.

Cyber experts Nadiya Kostyuk and Yuri M. Zhukov conducted a detailed quantitative analysis—with a data set cataloging 1,841 cyberattacks that took place between 2014–17—revealing that some 75 percent of the attacks between Russia and Ukraine originated from pro-Kiev forces.[119] Their research set out to test the effectiveness of cyberattacks as tools of coercion in war and concludes that cyber warfare did not have a measurable influence on kinetic

116. Yaroslava Denkovich, "The Verkhovna Rada Supported the Draft Law on Ukrenergo's Certification in the First Reading," *Kosatka Media Electricity News* (website), February 19, 2021, https://kosatka.media/en/category/elektroenergiya/news/verhovnaya-rada-podderzhala-v-pervom-chtenii-zakonoproekt-o-sertifikacii-ukrenergo.

117. Yaroslava Denkovich, "In 2020 Ukrenergo Completed Key Activities to Integrate the Ukrainian Energy System to ENTSO-E," *Kosatka Media Electricity News* (website), March 4, 2021, https://kosatka.media/en/category/elektroenergiya/news/v-2020-godu-ukrenergo-vypolnilo-klyuchevye-meropriyatiya-po-integracii-energosistemy-ukrainy-k-entso-e.

118. Ivan Nikitchenko, "Presumption of Guilt: How Microsoft Won a Protracted Battle on Unlicensed Software in Ukraine," IP Watch Dog (website), August 31, 2019, https://www.ipwatchdog.com/2019/08/31/presumption-guilt-microsoft-won-protracted-battle-unlicensed-software-ukraine/id=112810/.

119. Nadiya Kostyuk and Yuri M. Zhukov, "Invisible Digital Front: Can Cyber Attacks Shape Battlefield Events?" *Journal of Conflict Resolution* 63, no. 2 (February 2019): 325, https://doi.org/10.1177/0022002717737138.

military operations.[120] Even if cyberattacks have proven an ineffective tool in the conflict, Kostyuk and Zhukov provide clear evidence that pro-Kiev cyber forces are highly active.

In fact, Ukrainian hacktivism emerged early in the conflict, most notably in the groups comprising the Cyber Alliance, which targets the Kremlin as well as separatist forces. The Cyber Alliance argues that the leaks of sensitive e-mails, defacing of websites, denial-of-service attacks, and spear-phishing all contribute to weakening Russian efforts against Ukraine.[121] The government of Ukraine maintains its distance and intelligence officials are on record denying having any ties, though the organizations themselves assert that they receive limited support from the Ukrainian intelligence community.[122]

For their part, the pro-Kiev hacktivist forces are critical of the Ukrainian government's efforts to improve domestic cybersecurity. In 2018, Cyber Alliance controversially tested Ukraine's systems in an effort to find vulnerabilities. They reported 200 cases of vulnerabilities, including successful penetration of classified information in the Defense Ministry and, as previously mentioned, potential access to nuclear power plant systems.[123] Although admired by many, the Cyber Alliance poses a threat to more than the reputation of the Ukrainian government; their defiant stance against both Russia and their own state makes them a potential irritant in the growing relationships with Europe.

Not all cyberattacks originating from Ukraine are focused on disruption of Russian government or forces loyal to Russia. Ukraine was a notorious haven for cyber criminals prior to the war, known in particular for ransomware attacks. Early in the war, the United States tried to help the Ukrainian government reduce cyber crime and improve its poor record in bringing such criminals to justice.[124] Ukraine has since made significant progress. At the third US-Ukraine cyber dialogue in March 2020, the Federal Bureau of Investigation presented an award to Ukraine's National Cyber Police and the Prosecutor

120. Kostyuk and Zhukov, "Invisible Digital Front," 325.

121. Christopher Miller, "Inside the Ukrainian 'Hacktivist' Network Cyberbattling the Kremlin," Radio Free Europe/Radio Liberty (website), November 2, 2016, https://www.rferl.org/a/ukraine-hacktivist-network-cyberwar-on-kremlin/28091216.html.

122. Miller "Inside Ukrainian 'Hacktivist' Network."

123. Miller, "What's Ukraine Doing."

124. Mark Clayton, "How Ukraine Crisis Could Dent Country's Booming Cyber-crime," *Christian Science Monitor* (website) March 26, 2014, https://www.csmonitor.com/World/Passcode/2014/0326/How-Ukraine-crisis-could-dent-country-s-booming-cyber-crime.

General's Office for their efforts to arrest and extradite cybercriminals in conjunction with the Bureau.[125] In a joint operation in June 2021, the Ukrainian police arrested multiple suspects linked to the Clop ransomware gang, accusing them of responsibility for damages of about $500 million.[126]

Some of Ukraine's efforts to rein in cybercrime, however, have been controversial. In October 2020, in response to growing cyber incidents during the pandemic, draft laws increased police search powers and required Internet firms to provide access to large amounts of user data. These powers would make investigating cyberattacks easier, but, according to detractors, would endanger personal data in a country where corruption remains a critical issue.[127] The proposed law, however, is in compliance with EU standards, which is a top priority for Ukraine. As Ukraine tries to move closer to the EU and the ENTSO-E grid, harmonization of its legislation and credible enforcement of it will be important indicators of Ukraine's likely success.

Conclusion

The Ukraine case suggests several key lessons in energy security and CISR for NATO, its member states, and its partners.

- Attacks against critical infrastructure have been committed with state sponsorship. Even given this knowledge, the need for responsible attribution may delay a strong response.
- Cyberattacks have increased in their boldness, sometimes with unanticipated effects.
- The extreme networkness of electricity grids requires that parties to shared grids agree to and impose CISR standards.
- The intersection of plausible deniability and incomplete control of cyber actors can pose a danger both to the target state of cyber activity and to the state that hosts them.

125. Office of the Spokesperson, "Third Cyber Dialogue."

126. Carly Page, "Ukrainian Police Arrest Multiple Clop Ransomware Gang Suspects," *TechCrunch* (website), June 16, 2021, https://techcrunch.com/2021/06/16/ukrainian-police-arrest-multiple-clop-ransomware-gang-suspects/.

127. Umberto Bacchi, "Ukraine Plan to Tackle Hackers Sparks Privacy Fears," Reuters (website), October 7, 2020, https://www.reuters.com/article/us-ukraine-lawmaking-cyber-analysis-trfn/ukraine-plan-to-tackle-hackers-sparks-privacy-fears-idUSKBN26S1GG.

NATO defines cyberterrorism as a "cyberattack using or exploiting computer or communication networks to cause sufficient destruction or disruption to generate fear or to intimidate a society into an ideological goal."[128] The designation fits somewhat uncomfortably with the attacks in Ukraine, which took place in the context of a ground war between Russia and Ukraine. As this chapter has demonstrated, in this hybrid conflict, both states have tended to use cyber actors whose affiliation with the state is unknown or incomplete. Such a relationship is advantageous to the attacking state when it provides plausible deniability. The relationship becomes problematic when the attackers overreach or miscalculate the effects, potentially creating a dangerous escalatory cycle. The Russia-Ukraine cyberattacks reviewed in this chapter illustrate these complex balances. It is not clear that NATO or the EU can tolerate Ukraine's continued tacit support of cyber activists, no matter how much sympathy they may have for the cause. If Ukraine is to become part of the highly networked electricity system of Europe, rule of law must extend to cyber activist groups as well as to criminal groups.

Given the ongoing conflict with Russia, CISR in Ukraine is threatened by its proximity to Russia, by its shared energy infrastructure (pipelines and grids), by frequent reliance on Soviet legacy technologies, and by the normalization of cyberattacks in both directions. Through development assistance and shared practices and standards, the EU, NATO, and the United States are attempting to reduce Ukraine's vulnerabilities. Some progress is evident, but persistent problems remain. Infrastructure improvements are expensive and slow. Ukraine's cyber defenses continue to be hampered, according to Ukrainian critics, by "poor communication between state institutions, a resistance to change, a confused policy approach to cyber defense, and a lack of funds to recruit skilled personnel and buy much-needed equipment."[129] This criticism could be levied at many states, including the United States, especially in the wake of the May 2021 Colonial Pipeline attacks. Given the elevated threat environment in Ukraine, however, it remains an appropriate focus of US, NATO, and EU cybersecurity assistance.

Although Kostyuk and Zhukov find no evidence that cyber warfare had a measurable influence on kinetic military operations, the reverse appears true. Careful examination of escalation of cyberattacks appears to be a useful signal of coming escalation on the battlefield or border. Russia continues

128. NATO Centre of Excellence – Defence Against Terrorism, Responses to Cyber Terrorism (Amsterdam: IOS Press, 2008), 119.

129. Miller, "What's Ukraine Doing."

to pair cyberattacks with provocative military actions, as evidenced in April 2021. While Russian forces were massing on the border in April, there was a notable uptick in hacks and cyber aggression. *US News & World Report* reported that Ukraine, together with US partners, foiled 350 cyberattacks in March–April 2021 alone. This number is particularly striking given that Ukrainian intelligence reported only 600 total attacks in the previous year.[130]

Since energy has played an important role in the grievances between Ukraine and Russia, and has been used by both sides as a means in the conflict, it continues to be a likely target for cyberattack in the future, for as long as Russia pursues a hybrid approach. Strengthening critical energy infrastructure remains an important component of Ukraine's security, and increasingly of the EU's security as well. In the emerging era of electricity dominance, the vulnerabilities have shifted. Supply chains still matter, as they did in the eras of oil and gas, but the information chains within and between allies now constitute a critical vulnerability. Security of energy supply depends on improving the infrastructure and the oversight of not only EU member states, but also of its close partners. The Ukraine case provides a sense of the magnitude of the task and the potential way forward. Taking liberties with Churchill's observation about oil, it might be said that security now lies in allied cooperation, and cooperation alone.

Epilogue

This case study was completed months before the Russian invasion of Ukraine in February 2022. The book went to the press early in the conflict, but within the first two weeks it was clear that important events related to the cyber domain and the security and resilience of critical infrastructure were unfolding. The cyber element of Russia's war in the initial weeks was less aggressive than expected, which *New York Times* cyber experts posited was either due to Ukraine having better cyber defenses than Russia anticipated or to Russia's command decision to spare infrastructure in order to make

130. Paul D. Shinkman, "Russia Ramps Up Cyberattacks in Ukraine amid Fears of War," *US News & World Report*, April 20, 2021, https://www.usnews.com/news/world-report/articles/2021-04-20/us-helping-ukraine-foil-russian-cyberattacks-as-hacking-spikes-sources.

potential rule by a puppet government in Ukraine easier.[131] There is evidence to support both arguments.

Regarding the first hypothesis, the EU provided Ukraine a cyber defense team that officially constituted on February 22. Led by Lithuania, composed of members from Croatia, Estonia, the Netherlands, Poland, and Romania, and sponsored by the EU's Permanent Structured Cooperation defense and security initiative, the team had the collective mission to "detect, recognize, and mitigate cyber threats."[132] The EU's decision to employ this team was, in part, a response to the rejection of Ukraine's application in late 2021 to join NATO's Cooperative Cyber Defence Centre of Excellence as a contributing participant, which would have required a unanimous vote to succeed. Despite this formal rejection, Allies at the cyber center promised to provide ongoing support to Ukraine.[133]

Russia did launch some cyberattacks in the days leading up to its invasion of Ukraine, with analysts reporting significant cyberattacks against Ukraine on January 14, February 15, and February 23. The first two attacks were distributed denial-of-services (DDoS) attacks that targeted government ministries, banks, and defense agencies. The third attack, which took the form of a malware release, was potentially the most destructive. Microsoft's Threat Center in the United States, however, rapidly detected the new malware—known as FoxBlade—and blocked its code, provided this information with the Ukrainian government and, at the request of the US National Security Council, shared it with several Allies, including the Baltic states and Poland.[134]

In response, the Ukrainian government also mobilized its own "IT Army," which claimed to have more than 175,000 volunteers within days of being created on February 26. Prior to the IT Army's establishment, independent Ukrainian hackers claimed to have executed successful DDoS attacks against Russian targets and theft of data from Tetraedr, the Belarusian weapons

131. David E. Sanger, Julian E. Barnes, and Kate Conger, "As Tanks Rolled into Ukraine, So Did Malware. Then Microsoft Entered the War," *New York Times* (website), February 28, 2022, https://www.nytimes.com/2022/02/28/us/politics/ukraine-russia-microsoft.html.

132. Sebastian Sprenger, "European Union Cyber Defense Team Deploys to Aid Ukraine," *Defense News* (website), February 22, 2022, https://www.defensenews.com/global/europe/2022/02/22/european-union-cyber-defense-team-deploys-to-aid-ukraine/.

133. Sebastian Sprenger, "Ukraine Seeks Closer Ties with NATO on Cyber Defense," *Defense News* (website), February 1, 2022, https://www.defensenews.com/global/europe/2022/02/01/ukraine-seeks-closer-ties-with-nato-on-cyber-defense/.

134. Sanger et al., "As Tanks Rolled into Ukraine," (2022).

manufacturer.[135] Experts suggested that this group would undertake DDoS and defensive tasks to free up Ukraine's IT Army of hackers to go on the offensive against Russian targets.[136]

Although Russia conducted limited cyberattacks in the first days of the war, its direct attacks on Ukrainian critical infrastructure were unprecedented. Russian troops seized control of Ukraine's nuclear power plants at Chernobyl on February 25 and at Zaporizhzhya on March 4. Russian army kinetic attacks on these nuclear facilities prompted immediate international condemnation. Additionally, the IAEA accused Russia of undermining two of the seven critical pillars of nuclear power plant security: that operating staff must be able to fulfill their duties and make decisions, and that there must be reliable communications between nuclear power plants, the regulator, and others.[137] Although the conditions for the plant operators were quite unfavorable, Russian forces allowed them to continue management of the Zaporizhzhya and Chernobyl facilities. The Russian military command made no initial efforts to close the reactors at Zaporizhzhya, two of which were working at or near full capacity.

Russia's choice to allow the reactors to continue operations is particularly surprising given the fact that Ukraine had successfully de-linked from the Russian grid in its pursuit of joining the ENTSO-E grid. Ukraine had previously scheduled the isolation test—the required de-linking from the Russian electricity system to demonstrate that the Ukrainian grid could operate reliably in a stand-alone mode—for February 24, the same day that Russia initiated its invasion. Although it succeeded, the test was only supposed to last for a few days, at which point Europe would integrate Ukraine into the ENTSO-E grid. At the time of this writing, however, the members of ENTSO-E were still deciding whether to connect Ukraine directly with the European grid, effectively isolating Ukraine.[138]

Although it is possible for Russia to bring down the isolated Ukrainian grid without much difficulty, it has yet to do so. It is therefore not clear whether Russia's seizure of the nuclear power plants is part of a greater

135. Matt Burgess, "Ukraine's Volunteer 'IT Army' Is Hacking in Uncharted Territory," Wired (website), February 27, 2022, https://www.wired.com/story/ukraine-it-army-russia-war-cyberattacks-ddos/.

136. Burgess, "Ukraine's Volunteer 'IT Army.'"

137. "Update 13—IAEA Director General Statement on the Situation in Ukraine," International Atomic Energy Agency (website), March 6, 2022, https://www.iaea.org/newscenter/pressreleases/update-13-iaea-director-general-statement-on-situation-in-ukraine.

138. Suriya Jayanti, "Ukraine's Electrical Grid Shows How Hard It Is to Escape from Russia's Grasp," *Time* (website), March 1, 2022, https://time.com/6153039/ukraines-electricity-grid-escape-russia/.

plan to bring down the Ukrainian grid.[139] It may be, alternatively, an effort to ensure that command and control around the power plants will remain strong and avoid collateral damage to energy facilities by Russian troops, which was a well-documented occurrence in the Chechen wars. Given Russia's determination to conquer and control Ukraine, it may be that the power plants, especially the nuclear power plants, are not as attractive as targets in an open war as they had been in an era of hybrid conflict. Russia's seizure of these critical infrastructures may reflect its desire to retain control and security of these facilities in order to govern the territory in the future.

139. Jayanti, "Ukraine's Electrical Grid."

— 6 —

Civil Aviation

David Harell

Nearly every time civil aviation has been significantly challenged by terrorism since 9/11, the aviation security response, or lack of it, has not been successful in detecting and preventing the attack. In most of these cases, the anti-terrorism failure was exacerbated by intelligence and/or counterterrorism (CT) failures. In fact, the last time an attempt to blow up an aircraft midair was detected and thwarted by airport or airline screening procedures was the Anne-Marie Murphy incident at Heathrow Airport in April 1986.[1]

Terrorists have used multiple attack methods against civil aviation targets, including: hijacking; smuggling an improvised explosive device (IED) onto an aircraft by a passenger, by an airport employee (insider threat), or by concealing it in the cargo or a catering trolley; various types of sabotage; armed assault at an airport; suicide bombers at airports and on planes; man-portable air defense systems (MANPADS); and detonating an IED at an airport. Furthermore, there are additional attack methods that have been either carried out against other non-aviation targets or have been operationally considered by terrorists, and thus need to be addressed. Such examples include ramming into crowds with a vehicle, drone attacks with an IED, or drone swarm attacks against an aircraft while landing or taking off. Attacks can also include more than one modus operandi.

1. Philip Baum, "Ann-Marie Murphy and the Hindawi Affair: A 30th Anniversary Review," Aviation Security International (website), April 13, 2016, https://www.asi-mag.com/ann-marie-murphy-hindawi-affair-30th-anniversary-review/.

To discuss all these means and methods of attack as well as the potential mitigating measures would not be feasible in one chapter.

This chapter, therefore, will analyze the aviation infrastructure and the threats it faces, with primary focus on aircraft bombings and ground attacks on airports. To understand the civil aviation sector, the first section describes why the aviation industry is so critical, what makes it so volatile, and why it is so attractive a target to terrorists. The second section highlights several key reasons for the industry's vulnerability. The third section uses multiple case studies—which span the 20 years after the 9/11 attacks—to illustrate some of these vulnerabilities, examine the aviation security responses to terrorist attacks, and identify important lessons to be learned. Finally, the chapter concludes with recommendations and best practices that can assist in reducing the vulnerabilities across international civil aviation.

Understanding the Civil Aviation Industry

National and Global Critical Infrastructure

Civil aviation is defined by most countries as critical infrastructure due to the magnitude of the impact aviation has on both the global economy and the economies of individual nations around the world. Aviation plays a key role in multiple industries' supply chains, facilitates both tourism and business, enables disaster and pandemic response, and enhances the connectivity of the growing global network. See chapter 1 for discussion of transportation, and thus civil aviation, as a lifeline sector. The most recent example of its criticality is the role aviation has played in distributing the COVID-19 vaccinations globally. Although aviation cargo accounts for less than one percent of the quantity of goods shipped by the global supply chain, the monetary value of that percentage is 24 percent of the total value of all goods transported globally by all modes of transportation. The figures from 2019 demonstrate the scale of the aviation global infrastructure and illustrate its importance: (1) 4.3 billion passengers who traveled; (2) 38 million scheduled flights; (3) an economic impact of some $2.7 trillion; (4) 61.3 million tons of air freight transported; and (5) 65.5 million jobs to needed to sustain global air travel.[2] While civilian aviation is not inherently a military domain, it is clear that enhancing the security and resilience of this lifeline sector is a vital interest for NATO as an international organization as well as for its member states and partners.

2. "Slower but Steady Growth in 2019," International Air Transport Association (website), February 6, 2020, https://www.iata.org/en/pressroom/pr/2020-02-06-01/.

A Volatile Industry

Another major characteristic of the aviation industry is its volatility. The volatility relates both to economic factors and external geopolitical factors. The aviation industry is impacted by fluctuating demand, a rigid cost structure, competitive pricing, and changing and erratic fuel costs. In addition, the industry can be severely impacted by global events, such as the 9/11 terrorist attacks, the SARS and COVID-19 pandemics, and the global economic recession in 2008. These factors lead to an industry with low profit margins which can cause airlines to move from profitability to loss in a very short space of time. For example, in 2019, prior to the COVID-19 outbreak, profit margins for most US carriers were between 5 and 6 percent, which is generally considered low and susceptible.

An Attractive Target

In addition to the industry's volatility, it has become and continues to be an attractive target for terrorist attacks since the late 1960s. Since 1963, there have been more than 1,200 terrorist attacks against civil aircraft and airports, which figure 6-1 depicts by attack type over given periods of roughly 10 years.[3] This significant number of attacks demonstrates how attractive the aviation industry is for terrorists.

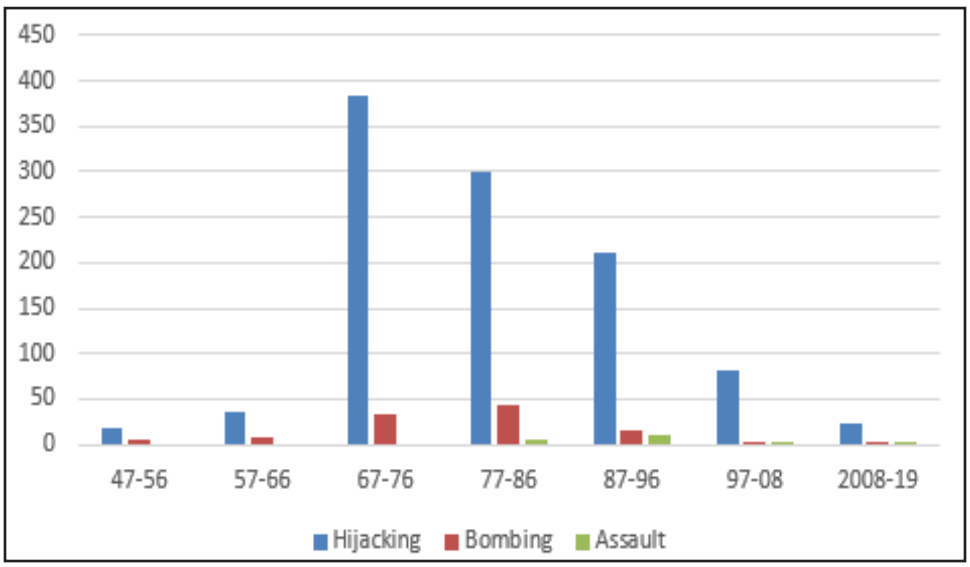


Figure 6-1. Distribution of terrorist attack types against aviation
(Data by Aviation Safety Network)

3. See "Aviation Safety Database," Aviation Safety Network (website), accessed September 28, 2021, http://aviation-safety.net/.

There are multiple reasons why terrorists see aviation as such an attractive target, including:

- The ability to achieve a mass casualty event. The majority of terrorists seek to carry out attacks that will achieve mass casualties. The higher the number of casualties, the greater the impact of the attack. By blowing up a plane in flight, there will likely be no survivors and large, modern planes can carry loads exceeding 400 passengers.
- The ability to cause significant damage to the economy of an attacked nation. In addition to the direct consequences of lives lost, there are significant indirect costs, particularly economic costs. A mass casualty terrorist attack against aviation can cause a drop in demand for air travel, with its negative impact on the volatile aviation industry.
- The intentional downing of an aircraft can be used as *casus belli*, which the terrorists may want to provoke. A terrorist attack against a country's aviation interests with the resulting dire consequences has in many cases been considered *casus belli*. Israel nearly went to war against Syria following the attempt to bomb an EL AL plane departing from London in 1986. The United States went to war in Afghanistan against al-Qaeda following the attacks of 9/11. In some cases, it can be the terrorist group's agenda to achieve an escalation of hostilities to further its cause.
- The psychological factor in terrorism. Terrorism is about achieving political objectives by terrorizing the public and thus putting pressure on governments. The fear of not being able to board a flight and reach a destination safely—but rather being blown up midair without any chance of survival—has a tremendous psychological impact on the public and the industry.
- National carriers are national symbols. National air carriers and other iconic airlines are seen by the terrorists as symbols of their respective nations. An attack against a national symbol is sometimes a preferred objective for terrorist groups.

- Immediate media coverage. Due to the centrality of the aviation industry to everyday life in the Western world, any event that has a negative impact on the industry receives immediate media coverage. This coverage may be disproportionate compared to other sectors. For example, a security event that causes the evacuation of an airport receives far greater coverage than a comparable event in another sector. Terrorist organizations are aware of this dynamic and media coverage is certainly one of their objectives.
- The adversary has certain operational advantages. Although not an exhaustive list, attacking an aircraft in flight offers adversaries several advantages, such as: (1) an extremely sensitive environment onboard a pressurized aircraft, which leaves little margin for error regarding the response and possible outcome; (2) the relatively small amount of explosives or limited weaponry required for an attack to have catastrophic consequences; and (3) no possibility for the aircraft to receive reinforcements in a hijacking situation.

The Aviation Industry Remains Vulnerable

Almost every time the aviation security (AVSEC) domain has faced terrorist attacks since 9/11, it has not performed well. In most cases the poor response was exacerbated by intelligence and/or CT failures. These failures clearly demonstrate that the aviation industry remains vulnerable. In addition, there have also been major security breaches at airports without a nexus to terrorism that add further concern regarding the vulnerability of the industry.

In table 6-1, the most significant attacks against civil aviation are listed from 9/11 to the present with a brief description of the nature of the AVSEC response and an indication of whether intelligence and/or a CT operation was able to provide any warning or disruption of the attack. The chapter will later expand on several of the attacks listed in the chart (those marked with an asterisk*), but not all of them; therefore, readers are encouraged to research these examples in more detail.

Table 6-1. Attack plots, AVSEC measures, and intelligence/CT inputs

Security Response		
Attack Plot	**Anti-terrorism/ AVSEC Measures**	**Intelligence/CT**
9/11	▪ Policy failure ▪ Catastrophic system failure	▪ Policy failures ▪ Failure to "join the dots" ▪ Failure of "imagination"
***Richard Reid "Shoe Bomber" December 22, 2001**	Failure: concealed IED not detected on two occasions and IED smuggled onto aircraft	FAA issues warning to airports and airlines on December 11 regarding the possibility that terrorists might put weapons in their shoes
Mombasa, Kenya al-Qaeda Combined MANPADS Attack and Suicide Bombings November 28, 2002	Security forces did not identify the cell launching the two MANPADS, which narrowly missed the aircraft due to terrorist error relating to missile's proximity to target	Lack of intelligence despite large al-Qaeda operation, including MANPADS weapon smuggling
***Thwarted Liquids Plot August 9, 2006**	Although airport security was not tested, the fact that a liquids ban was implemented after the plot demonstrates that airport security would not have detected the IEDs	Plot disrupted by CT operation
***Glasgow Airport Vehicle-borne IED (VBIED) Attack June 30, 2007**	No curbside antiramming measures in place, enabling VBIED driven by suicide bomber to crash into the terminal facade and deflagrate	Driver of the VBIED was on the MI5 terrorist watch list
***Northwest Airlines Flight 253 over Detroit "Underwear Bomber" December 25, 2009**	▪ Regular airport screening failed to detect IED concealed on Abdulmutallab's body ▪ Northwest Airlines security procedures at Schipol Airport fail to detect IED threat	US Senate Intelligence Committee found 14 intelligence failures leading up to the attempted attack aboard
Cargo Plot October 29, 2010	Cargo security screening measures fail to detect the IEDs	Intelligence warning arrives after IEDs have flown on two flight legs
***Brussels Airport Airside Diamond Heist February 18, 2013**	Airport security measures fail to prevent, detect, or respond to the severe perimeter breach and armed robbery of diamonds from a civil aircraft on the airside	N/A
Downing of Russian Metrojet October 31, 2015	Failure to prevent IED being penetrated onto the aircraft	Lack of information

Security Response (continued)		
Attack Plot	**Anti-terrorism/ AVSEC Measures**	**Intelligence/CT**
Bombing Onboard Somali Daallo Airlines **February 2, 2016**	▪ Screening measures bypassed by insider threat ▪ Security agent provided attackers an IED in laptop after the screening	N/A
Brussels Airport Suicide Bombings **March 22, 2016**	Failure to detect bombers or prevent them entering airport during attack deployment	Failure to act adequately on information received
Istanbul Atatürk Airport Armed Assault with Suicide Bombings **June 28, 2016**	Failure of external checkpoint to prevent entry of bombers into the airport	N/A
Islamic State (Da'esh) Sydney Plane Plot **July 15, 2017**	▪ It is not known whether the concealed IED would have been detected by Sydney AVSEC screening ▪ The investigation revealed the explosives used for the IED reached Sydney on a cargo flight from Istanbul	Intelligence warning received 11 days after first attempt to penetrate IED onto plane aborted by terrorist

Upon further examination, table 6-1 illuminates three major characteristics of aviation security that contribute to the vulnerability of civil aviation: its rigidity, its predictability, and its difficulty in keeping up with evolving terrorist threats.

Aviation Security Is Rigid

Security systems are comprised of personnel, technologies, and the regulations and standing operating procedures that determine how these elements should operate together. These systems are usually large and complex, consisting of many hundreds of employees and a multitude of technologies. The inputs of these systems (passengers, luggage, and cargo) are of very high quantity and frequency. The AVSEC system has to operate in an environment involving multiple stakeholders and adapting to political and legal constraints while maintaining high levels of customer service and satisfaction. Like many systems, aviation security works according to fixed procedures with little space to react differently to an irregular passenger or incident.

Two case studies demonstrate this rigidity of the system. The first case is that of Richard Reid, later infamously named the "Shoe Bomber."

Reid arrived at Paris Charles de Gaulle Airport on December 22, 2001, to board American Airlines Flight 63 bound for Miami, Florida, but security agents detained him for questioning and a search of his possessions. According to witness testimony, the security and airline personnel were "troubled, indeed perplexed" by Reid's disheveled appearance and his emotionless, calm behavior despite being subjected to a thorough inspection.[4] He was defined a higher-threat passenger and turned over to the police. However, the system's fixed search procedures, including both the French police and the airport security screeners, did not enable them to detect the IED concealed in his shoes on two occasions. Consequently, Reid was able to board the plane the next day with his concealed IED and attempted to detonate it.

The second example that also demonstrates the rigidity of the system is the case known both as the "Underwear Bomber" or "Christmas Day Bomber." Shortly before noon on December 25, 2009, a 23-year old Nigerian national named Umar Farouk Abdulmutallab attempted to perpetrate a suicide bombing aboard Northwest Airlines Flight 253, which was traveling from Amsterdam's Schipol Airport to Detroit and carrying 278 passengers and 11 crew members. Abdulmutallab ignited a small explosive device concealed in his underwear onboard the Airbus 330 as it was making its descent. Fortunately, due to the type of materials used as the explosive in the IED, the device did not explode, but rather deflagrated and burned Abdulmutallab in the process. Although there had been multiple warning indicators that should have had him placed him on the Transportation Security Administration's (TSA) watch list—thus preventing the attack—there are several AVSEC indicators which could have been identified. For example, Abdulmutallab paid cash for a one-way ticket from Lagos to Detroit (via Amsterdam), traveled to the United States in wintertime without any checked luggage, requested a window seat over the wing, and, according to witnesses, acted nervously at the gate in Schiphol airport. Despite all these indicators, Abdulmutallab was not singled out for any elevated screening procedures.

These two case studies demonstrate how flexible and adaptive adversaries can be, continually revising their modi operandi while AVSEC measures remained relatively rigid. Similarly, aviation security also needs to be flexible and able to adapt in real time to developing threats.

4. "Case Study: Richard Reid—The Shoe Bomber," X-Ray Screener (website), accessed November 9, 2021, https://www.x-rayscreener.co.uk/profiling/case-study-richard-reid-the-shoe-bomber/.

Aviation Security Is Highly Predictable

The AVSEC system relies on technologies and processes that have been in place for more than 30 years, especially in the case of walk-through metal detector gates and dual source X-ray machines. Even the more advanced computed tomography scanning machines have been in service since 1995. Seasoned passengers typically understand why these scanners sound an alarm, what item they are wearing triggered the alarm, or which particular item in their carry-on baggage has attracted the screener's attention. If routine passengers understand these triggers, how much more do well-trained and determined terrorists—who have access to these detection technologies—understand the capabilities and vulnerabilities in the system and take advantage of them. There are global terrorism online forums, for example, that specifically discuss the vulnerabilities of aviation detection technologies like X-ray machines and ways to exploit them, as figure 6-2 shows.[5]

Figure 6-2. Online terrorist discussion on how to outsmart X-ray machines
(Image by *Möbius*)

It is important to understand that the more predictable a security system is, the less likely the adversary is to be deterred by it. On the positive side, more advanced computed tomography scanning machines are gradually being deployed in some airports for screening carry-on luggage in North America and Europe, but these are more expensive than regular X-ray technology and consequently are cost inhibitive for many countries. Body scanners, too, are a significant enhancement to checkpoints, but they are often subject

5. Retired Chief Superintendent Michael Cardash, "Assessment of Jihadist Discussion on Lessons Learned from IED Attacks on Aircraft," *Möbius*, July 2014.

to psychological sensitivities leading to political pushback, which often impedes their implementation.

Aviation Security Has Often Struggled to Keep Up with the Threat

Analysis of the development of the threat against civil aviation from the 1960s until the present is characterized by adaptive adversaries constantly seeking ways to circumvent AVSEC measures. For example, in the 1960s and early 1970s, hijackings were the most prevalent form of terrorist attack. This reality led to the implementation of basic security measures (such as walk-through metal detectors and the screening of passengers' hand luggage). To circumvent these anti-hijacking measures, terrorists then focused more on plane bombings. As a result of the increase of plane bombings by IEDs in passengers' checked baggage, screening was improved, which in turn led terrorists to look for other ways to bypass these new measures. Some examples of these new measures include the use of MANPADS attacks or smuggling an IED onto the plane by concealing it in the cargo storage area as opposed to in a passenger's luggage, as was the case in the 2006 cargo plot. The thwarted liquids plot in 2006, which the next section will examine in greater detail, is an excellent example of security failing to keep up with the evolving threat. As detection technologies used in aviation security have improved, terrorists have also adopted ways to overcome these new technologies by creating more advanced, sophisticated IEDs.

As a brief survey of these examples demonstrates the vulnerabilities in the aviation industry, the next section will examine several case studies in greater detail to describe aviation security's various responses to terrorist attacks and identify important lessons the AVSEC community must learn.

Case Studies: AVSEC Responses and Lessons to Learn

Thwarted Liquids Plot, United Kingdom (2006)

On the night of August 9, 2006, following a major international CT operation, British authorities arrested 24 men and charged them with plotting a series of midair bombing attacks on transatlantic flights using explosive liquids onboard by concealing them in energy drink bottles. The plot was disrupted with the arrest of the cell members who were based in London and High Wycombe in Buckinghamshire. According to CT officials, the plot was a multiple suicide bombing mission inspired by al-Qaeda. During the arrests, police found martyrdom videos that showed cell members bragging about the attack, and materials from covert operations in east London showed

the cell members preparing the liquid bombs (see figure 6-3 for illustration of how the liquid bombs were constructed).[6]

Figure 6-3. Depiction of how the liquid IEDs were constructed
(Image by *Daily Mail*)

During the trial, the Crown alleged that the British terrorist cell planned to smuggle the IEDs onto the aircraft and blow up at least seven airliners departing from Heathrow Airport en route to North America with more than 1,500 people onboard. Beyond the passengers onboard these flights, the number of potential casualties must take into account those on the ground who would have been killed as a result of these attacks—a point that will be discussed later in the insights section. Video footage of tests conducted by government scientists, which were played to the jury, showed the devices producing an explosion powerful enough to blow a hole in an aircraft fuselage.

In May 2012, significant intelligence was gleaned from documents found on a concealed disk in the possession of a 22-year-old Austrian named Maqsood Lodin, who was being questioned by police in Berlin after returning from Pakistan. One document, written by Rashid Rauf—a British al-Qaeda operative at the heart of the group's terror operations in the United Kingdom—shed significant new light on the plot to blow up transatlantic airliners departing from Heathrow Airport in 2006. In the document, Rauf wrote: "We then analysed the various machines that were used for checking baggage and persons at airports. We found it was very difficult to detect liquids

6. David Williams and Rebecca Camber, "Three Guilty in Liquid Bomb Terror Plot to Murder Hundreds on Transatlantic Flights," *Daily Mail* (website), July 8, 2010, https://www.dailymail.co.uk/news/article-1293128/Three-guilty-liquid-bomb-terror-plot-murder-hundreds-transatlantic-flights.html.

explosives. After analysis that it would be possible to take concentrated hydrogen peroxide onboard, the thought came to our mind: would it be possible to detonate the hydrogen aboard an airplane?"[7]

Immediately after the plot's disruption, the United States and Canada banned all liquids and gels from passengers' hand luggage while in the United Kingdom no hand luggage was allowed onboard except for a few essentials (such as travel documents, wallets, and baby food). The new measures caused massive flight delays and threw most of the aviation world into disarray, but, over time, the ban was implemented in other countries that allowed passengers to take onboard only small quantities of liquid or gel. Niki Tompkinson, the then director of transport security in the UK Department of Transport, praised the way the United States, the United Kingdom, and Canada cooperated and coordinated their rapid response to this new threat of liquid explosives.[8] These restrictions have largely remained in force until the present.

In 2009, the plot's ringleader, Abdul Ahmed Ali, and his two closest associates, Tanvir Hussain and Assad Serwar, were found guilty of conspiring to bomb at least seven airliners flying to destinations in the United States and Canada. The High Court judge imposed life sentences with 30 minimum prison terms of 32–40 years each, calling the plot "the most grave and wicked conspiracy ever proven within jurisdiction," and comparable only to the 9/11 attacks.[9]

Liquids Plot: Insights and Analyses

One of the most important issues AVSEC professionals need to understand is whether the plot would have been successful if there had not been an intelligence warning and subsequent disruption by the CT operation. While some have questioned whether the plot was viable, Assistant Commissioner Andy Hayman, the then head of specialist operations at Scotland Yard, wrote: "If the plotters had not been stopped, I believe they would have been successful."[10]

7. Nic Robertson, Paul Cruickshank and Tim Lister, "Document Shows Origins of 2006 Plot for Liquid Bombs on Planes," *CNN* (website), April 30, 2012, https://edition.cnn.com/2012/04/30/world/al-qaeda-documents/index.html.

8. Niki Tompkinson, "Securing Transport in a Rapidly Evolving Environment" (address, International Homeland Security and Resilience Conference, London, May 29, 2007).

9. John F. Burns, "Life Terms for Plot to Bomb Trans-Atlantic Flights from London," *New York Times* (website), September 14, 2009, https://www.nytimes.com/2009/09/15/world/europe/15london.html.

10. Richard Greenberg, Paul Cruickshank, and Chris Hansen, "Inside the Terror Plot that 'Rivaled 9/11,'" *NBC News* (website), September 15, 2008, https://www.nbcnews.com/id/wbna26726987.

It is highly likely that had the plot materialized and not been thwarted, the terrorists would have had little difficulty infiltrating the security screening system in place at the time. The fact that the new security measures (for example, the ban on liquids and gels) needed to be authorized rapidly and implemented clearly demonstrates that measures in place at the time were not able to detect liquid explosives or the type of sophisticated initiating device the terrorists planned to use. In addition, one of the cell members who worked as a security screener at Heathrow Airport briefed his fellow cell members on the security measures in place. The cell's intimate knowledge of existing security measures further increases the likelihood that the attack would have succeeded. Details revealed by Rauf's aforementioned document indicate that the cell carried out significant hostile reconnaissance and had access to AVSEC detection technologies and procedures, which enabled the cell members to select their preferred modus operandi. This is an important insight to understanding the adversaries' capabilities.

The court's estimation, based on transatlantic plane passenger loads, was that at least 1,500 could have been killed in the bombings. The number of fatalities, however, would have been significantly higher if the terrorists had chosen to detonate their IEDs while the aircraft were flying over densely populated, large cities. These attacks, then, had the potential to exceed the number of fatalities caused on 9/11.

Another important issue that needs examination is why liquid explosives were considered a new threat in August 2006 despite the fact some types of liquid explosives have been around for more than a century. Of course, not all liquid explosives are viable for terrorist use against aviation targets due to their sensitivity and difficulty to access. However, PLX—a liquid explosive invented during World War II—was used by North Korean intelligence agents to blow up Korean Air flight 858 midair off the coast of Thailand on November 29, 1987, killing all 115 passengers and the crew. Here, North Korea's objective was to destabilize the region in the period leading up to the 1988 Olympic Games in Seoul. The North Korean agents left the sophisticated IED—consisting of a small quantity of plastic explosives concealed in a Panasonic transistor radio, a timing mechanism and detonator, and a bottle of whiskey containing liquid explosives—in an overhead bin onboard the flight and then disembarked the plane during a transit stop. Despite this use of liquid explosives for the first time against civil aviation, no new measures were put in place regarding the detection of liquid explosives. The emphasis, instead, was placed

on implementing measures to prevent persons from surreptitiously leaving a flight at an intermediate point.[11]

On December 11, 1994, liquid explosives were once again utilized in an attack against a civilian aircraft. In this attack, the notorious al-Qaeda terrorist Ramzi Yousef—currently in custody for his central role in the first World Trade Center bombing in 1993—planted a bomb on Philippine Airlines (PAL) Flight 434, which detonated en route from Manila to Tokyo. The plane survived the bombing and was able to make an emergency landing in Japan, but one passenger was killed and 10 more injured in the explosion. Yousef used nitroglycerine as the explosives component of the IED, concealed the liquid in bottles of contact lens cleaner with cotton balls serving as a stabilizing agent, and initiated the IED using a digital watch, two nine-volt batteries and a lightbulb filament (see figure 6-4).[12]

Figure 6-4. Reconstruction of the IED used to bomb PAL Flight 434
(Image by Alchetron)

Investigators would later learn that the PAL 434 bombing was a test run for a far deadlier attack known as the Bojinka plot. In this plot, Yousef and additional operatives in his al-Qaeda cell planned to blow up 11 aircraft departing from Asian airports and flying to the United States. The modus operandi was again to place an IED on the targeted aircraft and disembark the aircraft during a transit stop. The IEDs to be used were similar to the one tested in the PAL 434 bombing, including the use of liquid explosives.[13]

11. Billie H. Vincent, *Bombers, Hijackers, Body Scanners, and Jihadists* (Bloomington, IN: Xlibris Corporation, 2012), 50–53.

12. "Philippine Airlines Flight 434," Alchetron (website), July 28, 2021, https://alchetron.com/Philippine-Airlines-Flight-434#The-bomb.

13. John Hatzadony, "Oplan Bojinka Revisited," *Transport Security International* (website), October 16, 2019, https://www.tsi-mag.com/oplan-bojinka-revisited-the-plot-and-its-legacy/.

Fortunately, Yousef's plot was uncovered in Manila due to operational errors his cell committed. Had the plot been successfully executed, the impact could have been catastrophic and of a similar magnitude to the 9/11 attacks.

Taking into consideration the examples outlined above, three aviation attack plots all used liquid explosives as a component of their IEDs, based on the adversary's knowledge that it would be extremely difficult for security screeners to detect this type of threat. It is important for security policymakers to understand why the intended usage of liquid explosives in the 2006 plot was considered a new threat, as noted previously. According to relevant, senior AVSEC officials, there are four main reasons that the threat was not adequately addressed prior to August 2006:

- Over-compartmentalization created a knowledge gap between those working in intelligence and those in aviation security, leading to a situation in which not enough AVSEC decisionmakers were sufficiently aware of the viability of the threat or against it.[14]
- A general acceptance that existing detection technologies lacked the capability to detect liquid explosives and that it would take several years to develop such a technology.[15]
- Many countries believed that the threats they faced in civil aviation were not significant and that terrorists were focused mainly on the United States and Israel.[16]
- Prior to 9/11, many of the large airlines—which had significant influence on AVSEC policies—were opposed to new security measures that could prove costly and negatively impact customer service and facilitation.[17]

These four reasons shed important light on risk management issues that impact the way this threat has been dealt with in aviation security. The first point deals with the need for AVSEC leaders to have direct access to intelligence relating to the nature of the threats facing civil aviation. In response to 9/11, many senior police and military officers moved into the

14. Jacques Duchesneau, former president of Canadian Air Transport Security Authority, interview with author, March 9, 2021.

15. Billie H. Vincent, author and former senior US Federal Aviation Administration security official, interview with author, March 3, 2021.

16. Jonathan Zimmerli, former security inspector, Swiss Federal Office of Transport, interview with author, February 4, 2021.

17. Vincent, interview with author.

AVSEC domain. These were high-caliber professional security personnel, but they had a little or no background in aviation security. As they led their agencies to higher professional levels to counter post-9/11 threats, it is understandable that many of these officials would have no knowledge of the type of explosives used in terrorist attacks against aviation a decade earlier. This is an example of "not knowing what we know"—a familiar problem in the intelligence and security world in which an organization possesses information that could have improved its decision-making processes, but, for a variety of reasons, this critical information was not known or available to those who could have acted on it.

The second reason, related to a lack of technological capability to counter specific threats, essentially ignored the risk posed by liquid explosives. The fact that there were no technologies available for detecting liquid explosives and that it would take considerable time and money to develop such technology indeed has to be taken into consideration. It is important, however, to consider other possible responses in the procedural domain that could provide a certain response to the threat of liquid explosives.

It is a legitimate risk management decision to consider a risk and then decide not to address it as long as the residual risk is understood. It is necessary, however, when assessing the level of threat to a country's aviation interests to look at aviation as an entire network with the potential for negative, cascading effects. To illustrate this third point, the case of the "Underwear Bomber" is instructive. One could assume that the Netherlands in 2009 had evaluated the threat to its aviation interests as low and thus adopted a less than robust security posture, but was then was surprised by Abdulmutallab's plot. Abdulmutallab, a Nigerian living in London, was given the task of attacking a US airliner and decided to attack a flight with a transit stop at Schiphol airport. Had the plot been successful, the Netherlands—though not the target of the attack—would have been negatively impacted in many ways. The Netherlands actually has a robust AVSEC system and Schiphol Airport is considered a leader in the implementation of new security technologies. The IED concealed in Abdulmutallab's underwear would have succeeded in getting through security checkpoints in most airports due to the fact that in 2009 very few body scanners had been deployed at airports.

Regarding the fourth point, the influence of airlines over policies governing aviation security, this certainly was a contributing factor prior to the events of 9/11, and it could well be that commercial or other interests negatively impacted security policy decisions. Take, for example, the TSA's 2013 decision to allow passengers to board planes with folding pocket knives

with a blade less than 2.36 inches. The TSA made this risk management decision to enhance screening and passenger facilitation based on the logic that a small pocketknife poses no threat to an airplane, since all aircraft have reinforced cockpit doors that remain locked during the flight and meal service usually includes metal cutlery that is potentially as dangerous. This decision, however, created considerable pushback from the flight crew unions, among others, and in the end, the TSA had to cancel its decision.[18]

Compliance or Threat-oriented Aviation Security Systems?

Due to the sheer magnitude of passenger movements through airports both nationally and globally, the AVSEC community has struggled in one of its key management challenges: to calibrate a security level that will enable the screening of these passengers in an effective manner and in a reasonable amount of time. If the security screening level is too high, it can negatively impact passenger throughput. On the other hand, if the screening levels are too low, then they will not prevent or deter terrorists from defeating the system. A country's level of screening, known as the aviation security standard, is usually determined by the nation's AVSEC regulator. The International Civil Aviation Organization (ICAO), a specialized agency of the United Nations, serves as the global forum for international civil aviation and determines the international standard for aviation security. ICAO's most significant role is to develop and adopt standards and recommended practices for international civil aviation, which are incorporated into Annex 17 of the Convention on International Civil Aviation, also known as the Chicago Convention. ICAO's efforts to prevent and defeat illegal interference against civil aviation throughout the world is essential to aviation security and the future of civil aviation.[19] As the international standard with which all UN member states must comply, Annex 17—first adopted in March 1974—is updated on a regular basis, especially after a terror attack, a major security breach, or a significant rise in the threat level.

The large disparity in the resources that nations are able to invest in aviation security leads to significant differences in the level of detection technologies acquired and the availability of human resources and training on a country-by-country basis. Annex 17 provides an international standard which is quite general in nature and can be considered the lowest common

18. Andrew Bender, "TSA Cancels Decision Allowing Knives on Planes" *Forbes* (website), June 6, 2013, https://www.forbes.com/sites/andrewbender/2013/06/06/tsa-cancels-decision-allowing-knives-on-planes/?sh=4d592365e55a.

19. "Annex 17," International Civil Aviation Organization (website), accessed June 25, 2021, https://www.icao.int/Security/SFP/Pages/Annex17.aspx.

denominator for aviation screening. Therefore, many states have determined their own national standards for aviation security screening. For example, the TSA determines the US standard for aviation security as does the European Civil Aviation Conference for the European Union and Transport Canada for Canada. These three national aviation standards are more robust and stringent than ICAO's national standard, but even these advanced levels of security have not always been capable of dealing with adversaries' capabilities, as this chapter's case studies demonstrate. AVSEC officers work hard and invest considerable resources to comply with security standards. It is, however, important to note that being compliant with the standard does not necessarily mean being able to prevent, identify, and defeat a threat. There are many examples in which airports receive high compliance scores in their audits only to find that their security measures could not prevent an attack from occurring.

One example occurred at 8:00 p.m. on February 18, 2013, at the Brussels Airport, when eight masked gunmen in two vans cut through the airside perimeter fence and drove their vehicles onto the tarmac. In this very daring heist, the vehicles approached a Fokker 100 aircraft destined for Switzerland and the thieves, brandishing their weapons, held up the plane and stole $50 million worth of diamonds. The gunmen were then able to flee from the airport without being stopped and, fortunately, no one was hurt in this incident. The breach, however, must be considered from a different perspective: what if these gunmen were not thieves, but terrorists? The fact that eight well-armed gunmen were able to enter the airside operations area undetected highlights a very significant vulnerability in the airport's defenses. Here, the main point of failure was the airport's perimeter fence: since it was not a "smart fence"—equipped with sensors to detect intrusion and provide warning of a breach—the perpetrators were able to penetrate the perimeter without being detected. If there is no detection, then it is less likely there will be any effective response. It is important to note that neither the international standard nor the AVSEC standard in the EU requires airports to install smart fences.

The airport's spokesman insisted that security was entirely up to international standards, but highlighted that the heist was outside the scope of modern aviation security. In this case, security measures intended to deter or prevent "would-be bombers and other threats could not prevent commando-style raids by heavily armed criminals."[20] This incident is a classic

20. Andrew Higgins, "Brazen Jewel Robbery at Brussels Airport Nets $50 Million in Diamonds," *New York Times* (website), February 19, 2013, https://www.nytimes.com/2013/02/20/world/europe/thieves-steal-millions-in-diamonds-at-brussels-airport.html.

example of an airport being compliant with AVSEC standards but not being threat-oriented. An airport that decides to correctly invest in a smart fence is an airport going beyond compliance and seeking to be more capable to deal with the threats it faces. It is worth noting that this is the same airport that was attacked by suicide bombers three years later in 2016.

Need for Improved Physical Security Measures in Airport Public Areas

While the airside, checkpoints, international arrivals hall, passport control, and customs area of the airport are highly regulated and subject to regular audits, the airport's public areas—including the curbside, check-in terminals, and departure halls—do not receive the same regulatory attention. Consequently, the security levels in these public areas are usually determined by the airports themselves or by the airport police unit assigned to protect them, which can lead to significant disparity in security levels between airports even in the same country. When there is lack of regulatory policy or guidance regarding required security levels for the public areas, critical issues like stand-off, anti-ramming barriers, facade protection, and minimum armed security deployment often do not get properly or professionally addressed.

Terrorist ground attacks against airports since 9/11 have had significant impact on the airports attacked. One such attack occurred in the international departure hall at Brussels Airport on March 22, 2016, when two terrorists from the Islamic State (Da'esh) executed a double suicide bombing (a third bomber failed to detonate the IED in his bag and fled the airport). In addition to the 19 fatalities and over 80 injured, the massive destruction to the departure hall closed the airport completely for 12 days and only after months of partial operations did the airport resume full operations. Brussels Airport is Belgium's sole international airport, which made the impact of the closure even more substantial. According to multiple sources, the Brussels Airport and the city's metro system were under imminent threat, which prompted Belgian authorities to deploy military personnel to the airport and other key transport nodes throughout the city. The bombers were not detected by security personnel at the airport prior to the attack.

Similarly, the terror attack against Istanbul's Atatürk Airport in the evening of June 28, 2016, also involved three suicide bombers as in the Brussels airport attack, but in this case the terrorists from the Islamic State (Da'esh) were also armed with automatic weapons. Although this modus operandi of combining armed assault with a person-borne IED had occurred in the past—as in the case of the devastating attack against Bandaranaike International Airport in Sri Lanka in 2001—the Atatürk attack was the first time the tactic was used against a major Western airport.[21] The security deployment at the Istanbul Airport was quite different from the one in Brussels because Turkey had seen a significant upsurge of terrorism in the previous 12 months, including six significant bombing attacks carried out in the first six months of 2016. Turkish security authorities deployed screening points at the curbside entrances to the terminals. The deployment of curbside checkpoints can be a very effective countermeasure against suicide bombers because they signal a deterrence posture and enable identification of the threat before would-be bombers can enter the terminal.

As with all security measures, however, there are trade-offs. Curbside screening points can also create a bottleneck situation where people are stalled as they wait in line to pass through the security checkpoint and become the target of attack. The main drawback with curbside screening points is that they negatively impact facilitation and customer service and are not easily accepted by the key stakeholders: the airports and the airlines. In the Istanbul Airport attack, it is likely that the casualty numbers and damage to the airport would have been significantly higher had all three bombers successfully detonated their suicide bombs inside the terminal. In fact, the airport was able to resume flight operations the morning after the attacks, as opposed to the lengthy closure Brussels Airport experienced. The resilience demonstrated by airport operators in Istanbul should be commended: the police and security forces deployed at the airport responded bravely and quickly, which contributed to mitigating the threat and minimizing the number of casualties and the damage to the airport.

A final case study that deals with a terrorist ground attack against an airport involves the use of vehicle-borne IED (VBIED). On June 30, 2007, two terrorists drove a sport utility vehicle into the glass doors of the main terminal (Terminal One) of Glasgow International Airport. The slow speed of the vehicle prevented the vehicle from actually entering the crowded terminal. The two terrorists tried to detonate the VBIED, but fortunately

21. Rohan Gunaratna, "Intelligence Failure Exposed by Tamil Tiger Airport Attack," *Jane's Intelligence Review* 13, no. 9 (September 2001): 14.

it only burst into flames and resulted in a delayed, minor explosion. The only casualties in the attack were the two terrorists, one of whom died and one of whom was severely burned. Despite the relatively low impact of this attack, the "what if?" question is again worth asking. What if the VBIED had been better constructed and had detonated? What if the vehicle had succeeded in penetrating into the terminal building? Based on the destruction and number of casualties that person-borne IEDs caused in Brussels, it is likely that the VBIED—had it detonated inside the terminal building—would have exceeded that level of devastation and casualties considerably.

It is important to note that there were no anti-ramming barriers or bollards along the curbside at the Glasgow Airport prior to the attack. After the attack, however, bollards were installed along the curb to protect the terminal from attacks of this nature. See figures 5 and 6 for images of the airport curbside before and after the attack.[22] Although many airports around the world are now protected by anti-ramming barriers or bollards along the curbside, many others still do not have this protection. Again, this is an example of a lack in regulatory policies for the public areas in airports.

Figure 6-5. Glasgow Airport curbside before the 2007 VBIED attack
(Images by Insider)

22. Steven Wilson, "Glasgow Airport—Then and Now," *Insider* (website), June 30, 2017, https://www.insider.co.uk/news/gallery/glasgow-airport-then-and-now-10713904.

Figure 6-6. Glasgow Airport curbside after the 2007 VBIED attack
(Image by Insider)

Recommendations and Best Practices to Reduce Vulnerability

While the third section analyzed case studies of terror attacks against airports and aviation since 9/11 and discussed possible root causes that have influenced the outcomes of these attacks and plots, this final section presents key recommendations—based on experience and best practices—aimed at reducing the vulnerabilities identified throughout the chapter.

Develop a More Risk-based AVSEC Screening System

As the previous section discussed, one of the key risk management dilemmas in aviation security is finding the optimal screening level that ensures the necessary throughput speed and customer service without compromising the ability to prevent or deter terrorist attacks. On the one hand, a security screening level that prioritizes efficient access and minimizes congestion at airport checkpoints may not be able to detect or prevent a sophisticated attack against the system. On the other hand, enforcing a screening protocol that enables the detection of even the most sophisticated, concealed threats may negatively impact airport throughput in such a manner that airport operations will not be sustainable. Most nations choose to implement

a "one size fits all" AVSEC system in which everyone is screened in the same manner and at the same level.

Since many of the attacks discussed in this chapter were able to penetrate this type of AVSEC screening successfully, many regulators now promote a more risk-based approach. Two former administrators of the TSA, Kip Hawley and John Pistole, were two of the first advocates for moving toward a more risk-based model of aviation security.[23] A risk-based system means that passengers considered a higher threat will receive a higher level of screening while routine, lower threat passengers are screened according to the minimum standard. Here, physical screening levels can vary in accordance with a passenger's potential risk. An important complementary issue with this recommendation is the decision on which tools to use to identify passengers with a potential for higher risk. This issue will be addressed further in the additional recommendations of this section.

Develop and Implement Threat Definitions Aligned to Adversary Capabilities

The foundation for effective security systems in aviation, and all other security sectors, should be to define correctly the threat criteria that need to be addressed and then prepare the response accordingly. Security work inherently must relate to threats; without threats, there would be no need for security measures. In the aviation sector, if there were no terrorists or threat actors with malicious intentions against aviation assets, then there would be no need for aviation security. Furthermore, securing all aviation assets against every type of terrorist threat would demand tremendous resources, which—even if they were available—would almost certainly not be feasible or cost-effective. The direct consequence of attempting to counter all possible threats is that security personnel will not be adequately focused on the relevant threats and thus will provide only a partial solution. For example, continuing to define nail scissors or small pocketknives as a threat to aviation security makes it more likely that the efforts screeners undertake to detect these items will decrease their ability to detect other, more sophisticated and relevant threats.[24]

23. Hugo Martin, "TSA Head Envisions End of One Size Fits All Security Measures," *Los Angeles Times* (blog), March 3, 2012, https://latimesblogs.latimes.com/money_co/2011/03/tsa-head-envisions-end-of-one-size-fits-all-security-measures.html.

24. A. T. Biggs et al., "Examining Perceptual and Conceptual Set Biases in Multiple-target Visual Search," *Attention, Perception, & Psychophysics* 77 (2015): 844.

It is important that regulators provide clear definitions of the threat so that security managers and frontline staff know exactly what they are trying to detect and defeat. Vague threat definitions can lead to frontline staff not correctly calibrating their equipment, or, worse still, if a threat is not clearly defined, then it may not be addressed at all. If ramming into the terminal is not defined as a threat by the regulator, then airport operators may not feel obliged to install anti-ramming barriers. In short, threat definitions must be quantitative, qualitative, and scenario-based, such as an armed assault on the terminal by an adversary comprising of two to three terrorists armed with automatic weapons and IEDs.

Utilize Airline Passenger Travel Data for Risk-based Screening Purposes

Passenger name records (PNR) are commercial records for storing airline reservations and records related to other travel services. For example, a single PNR can contain data about a single traveler, an entire family or tour group, and all services for their trips from multiple providers, such as air and train travel, hotels, and car rental. PNR data can be most insightful because it reveals a passenger's links and connections, activities, tastes, and preferences. PNR data typically contains credit card numbers, telephone numbers, e-mail addresses, Internet protocol addresses, and place and mode of payment. A focused analysis of a passenger's PNR can provide important indications that something is amiss and/or irregular. These indications, sometimes termed *red flags*, do not always mean that there is illegal activity connected to them. History shows, however, that in almost all cases, post-attack examinations reveal that red flags were present.

At the time of the 9/11 attacks, the US Federal Bureau of Investigation and the Federal Aviation Agency regulated the airlines to use a computerized system that utilized both PNR data and a projected profile of a terrorist, known as the Computer Assisted Passenger Pre-screening System (CAPPS). The objective of the system was to identify potentially high-risk passengers whose bags would then be screened for explosives and not allowed to be loaded on the plane until the passenger had himself boarded. Unfortunately, the focus of the CAPPS system was a flagged passenger's checked baggage; the system had no impact on checkpoint screening levels for passengers and their carry-on bags. In the case of the 9/11 attacks, CAPPS identified 11 of the 19 hijackers on three out of the four planes on that fateful day but was unable to intervene

to prevent the attacks.[25] This failure is a good example of the AVSEC system being inflexible and rigid because it was able to detect but not adapt to the threat of suicidal hijackers. Experts in aviation security understand that PNR data is currently not being exploited to its full potential. PNR data contains powerful indicators that should be used to drive a risk-based approach to screening. In order to adopt this approach, legitimate privacy concerns must be addressed and international agreements put in place regarding the sharing and protection of PNR information.

Integrate Behavioral Detection Programs

In addition to utilizing PNR information, another important tool that can be used for risk-based screening is behavior detection or passenger observation programs. The concept of behavior detection works much the same way as using PNR data. While PNR data reveals indicators or red flags in a passenger's records, behavior detection focuses on passengers' behavior or their contextual circumstances. It is important to note that this practice is not racial profiling of any form but rather identifying suspicious behaviors and activities that fit the profile of aviation terrorists. Suspicious behavior indicators are the result of the terrorist being under stress, and they include agitation, aggressiveness, and other behaviors that differ from what bona fide passengers typically exhibit.

Several countries—including, but not limited to, Israel, Singapore, the United Kingdom, and the United States—have already implemented passenger observation programs at airports with notable success in the detection of criminal activities. ICAO has registered behavior detection in Annex 17 of the Chicago Convention as a recommended AVSEC practice for all countries to adopt.

Design and Implement Airport Community Security Programs

Another challenge in aviation security is how to enhance capabilities without depending on resource-intensive solutions, which are costly and, in most cases, unavailable. One way to achieve this goal is to involve the entire airport community in the security effort. Safety officers and their teams will not achieve a high level of safety without the involvement of all stakeholders in the facility, institution, or factory where they work. In other words, it has become accepted that safety needs to be everyone's business, and this concept has been highly successful. This same concept can be implemented

25. National Commission on Terrorist Attacks upon the United States, *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States* (Washington, DC: National Commission on Terrorist Attacks upon the United States, 2004), 1–3.

in the AVSEC arena with the same levels of success. The idea is to have all stakeholders at the airport, including non-security employees, contribute to the security effort. The objective is to effectively utilize non-security resources as a force multiplier for security. One may ask why these airport employees should contribute to the security effort in addition to their everyday job—and without additional payment. In short, it is in their best interest for the airport to be safe and secure, since a terrorist attack not only threatens their livelihoods but their lives as well. Therefore, with the right explanation and justification, it is possible to get most employees to support such a program.

Airport community security programs do not expect non-security employees to dedicate a large amount of time to security work, but when every employee contributes a small amount to security it bolsters the overall level of security. For example, the ticket agents who check in passengers for their flights are usually not involved in any of the airport's security processes, but these agents have more interaction with the passengers than most security personnel at the airport, and thus, they have the best knowledge of the types of passengers frequenting the different flights. With a short, focused training session, these ticket agents can be taught to identify irregular indicators or red flag behaviors, and these inputs can be incorporated into a risk-based screening system. Another example is the inclusion of airport employees working at information desks in the airport halls. These employees spend much of their time observing the activities inside the airport halls and waiting for passengers to request assistance. Again, with a little training, these employees can assist airport security by identifying and reporting irregular activities. This training can be done with all the different employee groups at the airport, to include vendors, cleaners, baggage handlers, and airport traffic wardens.

These types of community security programs are highly effective and significant force multipliers that improve local and international aviation security. To implement such a program requires buy-in from all the relevant stakeholders and also a central organization to lead, design and implement such a program. In Singapore, the government implemented such a program at Changi airport and other border crossings with significant success. The program—known as the threat-oriented person screening integrated system (TOPSIS)—achieved in its first year of implementation a 60 percent increase in the detection of criminal activity at the airport without any

interruption to throughput or passenger service.[26] TOPSIS has had many additional successes in the homeland security domain, which readers can further research online. Airport security community programs, like TOPSIS, represent a modern approach to security that can add significant resources to enhance airport security, while making it more flexible and less predictable to the adversary.

Harden Airport Perimeters

Although there have been no terrorist attacks on the airside of airports in recent years, there have been other criminal events that involved breaching the airport perimeter. In addition to the 2013 diamond heist at Brussels Airport discussed earlier in the chapter, two notable vehicular breaches occurred at the airports in Lyon, France, in 2018, and in Van Nuys, California, in 2020.[27] These incidents had no nexus to terrorism, but they clearly demonstrate the existing vulnerabilities of the airside and highlight the importance of hardening the perimeter to prevent possible terrorist attacks. Of course it is not the intention to turn the airport perimeter into a fortress. What is required is that any attempt to cut, climb, or ram through the perimeter will trigger an alarm that will allow airport security to respond and intercept the threat while it is at a distance from the airport terminals or taxiing aircraft. Some airports have already invested in robust perimeter fencing but it will take regulatory action to raise the required standard for all. The following section will further discuss the need for regulatory action as the next recommendation.

Improve Regulation of the Airport's Public Areas

As the section on ground attacks previously discussed, there are no or very minimal security requirements regarding physical security measures that should be implemented to protect the ground side or public areas of airports. There should be minimum standards relating to the structure of an airport, facades, curbside protection, unattended vehicles, and armed deployment. These standards would be especially relevant for new terminals in the predesign stage and for terminal upgrades. Implementing security measures

26. S. Iswaran (address, TOPSIS Forum 2011 Cum Counter-Terrorism Exhibition, Singapore, October 12, 2011), https://www.mynewsdesk.com/sg/ministry-of-home-affairs/pressreleases/topsis-forum-2011-cum-counterterrorism-exhibition-at-cag-auditorium-changi-airport-terminal-2-speech-by-mr-s-iswaran-minister-in-prime-693260.

27. Kim Willsher, "Dramatic Police Chase as Car Smashes onto Runway at Lyon Airport," *Guardian* (website), September 10, 2018, https://www.theguardian.com/world/2018/sep/10/french-dramatic-police-chase-as-car-smashes-onto-runway-at-lyon-airport; and CBSLA Staff, "Police Chase Down Erratic Driver on Van Nuys Airport Runway," *CBS News Los Angeles* (website), November 6, 2020, https://losangeles.cbslocal.com/2020/11/06/police-chase-erratic-driver-van-nuys-airport-runway/.

during the design stage of a construction project is cost-effective way to achieve the desired level of security with a minimal impact (low single-digit percentage) on the overall project cost.

Avoid Over-reliance on Indications and Warning Intelligence

Experience has shown most terror attacks against aviation have occurred either without warning or with intelligence that was too general to assist in focusing the AVSEC response. It would be prudent for AVSEC officials to hold the basic assumption that an attack against aviation interests can happen at any time and without any prior intelligence warning. It is also important that relevant intelligence related to aviation threats and terrorist capabilities be shared with the officials charged with protecting aviation assets. It is relatively easy to keep security personnel alert and prepared during a period of heightened threat, and it is therefore an important function of security leaders to use tools that will maintain a high level of alertness even when there is no intelligence warning.

Airport security departments should also develop their own field intelligence-gathering capabilities, with a focus on unusual events that take place at the airport, which could be tests of security or hostile concepts. These events should be logged and monitored for trends. Airport field intelligence should also be shared with the broader AVSEC community, which may enable the detection of patterns of activity.

Prioritize the Human Factor: Recruitment and Training

This chapter demonstrates that aviation security is a complex system that deals with sophisticated threats while utilizing modern technology. In order to meet the multiple challenges the industry faces, it is imperative that AVSEC personnel, both managers and frontline officers, are of the right quality and receive the correct training. For too many years, airport screeners were not adequately trained or rewarded, and in many cases the position of an airport screener was an entry-level job. This type of personnel management can ultimately lead to situations in which the entire system could fail or collapse due to a single frontline security screener's human error. The extraordinary event that took place at Edmonton International Airport on September 20, 2013, illustrates this point. Here, an airport screener stopped an 18-year-old passenger after identifying what looked like a pipe bomb in the passenger's carry-on bag. The passenger explained that he had constructed the item in the past as a hobby, and he had forgotten it in his bag. Unbelievably, the screener told the passenger he could keep the item onboard the plane. The passenger, however, insisted the item remain with the screener,

so the screener accepted the device and allowed the passenger to board his flight to Mexico. The device was six inches long and two inches in diameter and filled with gunpowder. It had screws at both sides and a three-meter fuse running through it. The pipe bomb was placed in a bin with other confiscated items until a supervisor notified the police at the airport four days later.[28]

To prevent events like this from recurring and to develop a professional, robust security posture, it is essential to recruit the right people and also to train and reward them. Threat-oriented exercising, or red teaming, is another essential and powerful tool for building and maintaining the professional readiness of AVSEC personnel and reducing the level of human error. Integrating technology is extremely important to aviation security and it will become increasingly so, but without professional, alert, and motivated officers, achieving a more secure, less vulnerable aviation industry will not be possible.

Conclusion

This chapter focused on the particular characteristics of the civil aviation sector as critical infrastructure on a national and global level. Aviation security is a complex system working in an increasingly challenging environment. Terrorists will continue to target aviation assets globally for the many reasons discussed in the chapter and seek to exploit the different vulnerabilities the various case studies revealed. Students and practitioners of security and risk management must understand there is no such thing as 100 percent security, yet the consequences of multiple, sophisticated attacks against civil aviation can be devastating in terms of loss of life and economic impact. Therefore, attaining a security level that mitigates these kinds of threats while enabling the aviation industry to carry out its operations and achieve its objectives is essential. The different recommendations discussed provide an opportunity to improve capabilities across aviation security. AVSEC professionals, and those who support and enable them, should be continually seeking ways to improve performance, reduce vulnerability, and increase deterrence against would-be attackers.

28. Charles Rusnell and Jennie Russell, "Edmonton Pipe Bomb: Airport Security Personnel Ignored Danger," *CBC News* (website), January 17, 2014, https://www.cbc.ca/news/canada/edmonton/edmonton-pipe-bomb-airport-security-personnel-ignored-danger-1.2500071.

— 7 —

Mass Transit Railway Operations

Adrian Dwyer

This chapter considers the vulnerability of railway operations to terrorist action. Its focus is on those methods of attack that have been used previously yet remain relevant. Highlighted throughout is the vulnerability inherent within open transport networks that are also tightly coupled systems. In such systems, a disruptive incident has the propensity to ripple quickly across the network and precipitate a range of unintended and often lethal consequences. From the perspective of the North Atlantic Treaty Organization (NATO), the targeting of rail networks across NATO member states can also disrupt military logistics, the civilian supply chain, and economic prosperity more generally. This chapter is not configured in the manner of a threat assessment and does not predict future events in time or space. Instead, it is arranged in four sections and uses plausibility as an appropriate leitmotif for managing terrorism-related risk under conditions of uncertainty.

- Section 1 addresses the inherent vulnerability of rail and the need for a proportionate approach when assessing risk in context. This section discusses how to counter the complexities of the threat in terms of addressing the perceived threat within a framework of plausibility. It outlines one approach to the process of strategic risk assessment.

- Section 2 examines the multifaceted nature of railway operations and further develops the scale and scope of the risk management challenge. It notes that policing and security are not interchangeable concepts, and that benefit, in terms of public safety and managing societal risk, is maximized when these endeavors operate seamlessly.
- Section 3 uses case study data from Great Britain, continental Europe, the United States, Japan, and India to contextualize how threat actors exploit inherent vulnerability. These 15 exemplars range from unsophisticated physical assaults to large-scale mass casualty events.
- Section 4 summarizes the lessons available. It notes that certain themes within methods of attack often recur and that this observation is particularly relevant when attempting to manage risk under conditions of uncertainty. The section concludes by noting that though the threat from terrorism is diverse, inherent vulnerability can be mitigated.

Railways Are Vulnerable by Design

Inherent Vulnerability and the Strategic Assessment of Risk

Mass transit and freight networks utilizing a permanent railway infrastructure are inherently vulnerable to terrorist attack. Here, vulnerability relates to open access, predictable and rigid operating parameters, and the target-rich nature of the operating environment. Also relevant is the strategic importance of the railway's socioeconomic functions. The nub of the counterterrorism (CT) challenge, therefore, is to protect passengers and rail assets while not compromising the primary purpose of the enterprise: to facilitate the unhindered movement of people and goods in a timely and efficient manner. Implicit within the term efficient are the constructs of safety and reliability.

To operate an efficient network, risk must be foreseen, prioritized, and managed appropriately.[1] As part of the CT challenge, establishing a scientific probability for the manifestation of a malicious act, within useful temporal and spatial parameters, is problematic.[2] There are numerous variables,

1. *Blackett Review of High Impact Low Probability Risks* (London: Government Office for Science, 2011), 7–8, https://www.gov.uk/government/publications/high-impact-low-probability-risks-blackett-review.

2. Brian Appleton, *Appleton Inquiry Report* (London: Health and Safety Executive, 1992), 4, 27–29.

some latent and others specified inadequately.[3] Under conditions of such uncertainty, any casual use of the term likelihood can be misleading and may be taken to represent a level of confidence that is unwarranted given the scale of the unknowns. In contrast, determining the plausibility of a particular method of attack (MoA) and its foreseeable impact in a defined context is a more inclusive process that accommodates inevitable uncertainties. Threat information is considered against reasoned judgment based upon evidence, relevant experience, special knowledge of the operational environment, and, by necessity, explicit assumption. Figure 7-1 illustrates the relationship between these terms in the context of determining risk.[4]

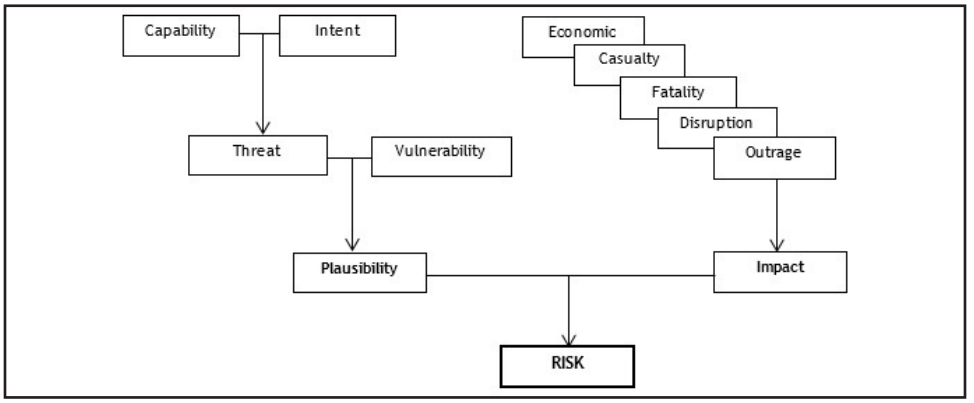


Figure 7-1. Relationship between terms in risk assessment
(Diagram adapted from the UK Cabinet Office)

Developing a strategic risk assessment (StRA) involves combining multiple risk assessments within a context-specific framework that considers plausible MoA in concert with the range of credible intentions ascribed to threat actors. This inclusive approach combining qualitative and quantitative data is underpinned by the concept of the reasonable worst case (RWC). In the United Kingdom, RWC is "the worst plausible manifestation of that particular risk (once highly unlikely variations have been discounted)." This form of words represents a more generalized use of language than that noted previously, where malicious acts were defined against variables

3. For additional information, see Chief Coroner, *Inquests Arising from the Deaths in the London Bridge and Borough Market Terror Attack* (London: Royal Courts of Justice, 2019), 18, https://londonbridgeinquests.independent.gov.uk/wp-content/uploads/2019/11/Final-Report-on-Action-to-Prevent-Future-Deaths-Report.pdf; *National Risk Register of Civil Emergencies* (London: Cabinet Office, 2017), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/644968/UK_National_Risk_Register_2017.pdf; Karen L. Petersen, "When Risk Meets Security," *Alternatives: Global, Local, Political* 33, no. 2 (2008): 185; and Nancy K. Hayden, "The Complexity of Terrorism: Social and Behavioral Understanding Trends for the Future," in *Mapping Terrorism Research*, ed. Magnus Ranstorp (London: Routledge, 2006), 304.

4. *National Risk Register* (2017), 69–71.

of plausibility and impact. Likelihood was applied only to accidents and natural hazards—that is to say, those events for which data were considered more bountiful and reliable.[5]

In assessing risk holistically, the quantified estimates of vulnerability and a target's perceived attractiveness to threat actors will be influenced by precommitments and biases of the raters, hence the requirement for a more inclusive and transparent process that addresses the following issues:

- Which risks are plausible, rather than everything imaginable
- The operational context in which the plausible risks are applicable, such as which constituencies they affect, how they may combine, or if they are spatially or temporally specific
- By what means the rationale supporting the ranking of risks is qualified

A situation in which risk assessments are not qualified within a strategic framework and are informed mainly by generalized intelligence material, and in which stakeholders adhere knowingly or unknowingly to differing worldviews, is less likely to produce the outcome desired. Such a situation also represents an approach to risk management that can be complex to justify in the face of hostile scrutiny. See also the useful commentary on risk assessment and management in chapter 13.

Rail incidents, because of the nature of the tightly coupled system in which they occur, have the potential to escalate rapidly in time and space, and this problem is exacerbated when terrorism is the known or suspected cause.[6] Raising the specter of terrorism places extra demands on often hard-pressed cognitive resources of decisionmakers. For example, does initial reporting represent the end of the attack phase or its beginning? Are terrorists still at the scene, waiting to exploit inevitable confusion and target responders? Are other attacks pending elsewhere on the network? These are all valid

5. *National Risk Register* (2017), 8–10; *National Risk Register*, 2020 edition (London: HMG, 2020), 8, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/952959/6.6920_CO_CCS_s_National_Risk_Register_2020_11-1-21-FINAL.pdf.

6. For additional information, see Donald Holbrook, Gilbert Ramsay, and Max Taylor, "'Terroristic Content' towards a Grading Scale," *Terrorism and Political Violence* 25, no. 2 (2013): 202–23, https://doi.org/10.1080/09546553.2011.653893; David Anderson, *Report on the Terrorism Acts in 2011* (London: Stationery Office, 2012), 26; Maura Conway, "The 'T' Word: A Review of Richard English's Terrorism: How to Respond," *Irish Literary Supplement* 30, no. S1 (2010): S1–S4; and Charles Perrow, *Normal Accidents* (Princeton, NJ: Princeton University Press, 1999), 89–92.

questions, but reliable answers may be unavailable, which further accentuates the effects of uncertainty on the decision-making process.[7]

This chapter adopts a case study methodology, using 15 unique exemplars, to provide an overview of the vulnerability of rail in context and highlight some resilience-led response options. The intention of this methodology is to communicate the essence of what happened in these situations in a recognizable form and to help develop foresight from the lessons of others. The approach accommodates variations noted in relation to how railways operate and how threat actors exploit resultant vulnerabilities. The exemplars focus on what history demonstrates are preferred targets: passengers and the vulnerable elements of accessible infrastructure, rather than goods in transit.[8] The historical focus of these preferred targets does not mean, however, that some types of freight movement—nuclear materials or certain types of chemicals, for instance—do not have a special attraction for threat actors. In considering the line of route (LoR), long-term denial through terroristic endeavor is rare. Deliberate or hasty demolitions, particularly of monolithic structures, require expertise and resources not possessed by many aspiring or even some well-established terrorists. This observation is central to establishing a rational appreciation of a threat actor's capability and intentions—that is to say, the plausible threat they are perceived to pose. See also the commentary on physical threats in chapter 2.

Absent from the range of exemplars in this chapter is cyberterrorism, which is an established area of growing concern. In terms of networked rail systems, cyberattacks have not featured as an effective MoA and, to date, cannot be equated with those incorporating physical means, such as improvised explosive devices (IED), firearms, and attacks involving sharp, bladed, or blunt (SBB) weapons.[9] Those cyberattacks recorded within the public domain are most typically denial of service-type activities, often related

7. Terje Aven and Ortwin Renn, "The Role of Quantitative Risk Assessments for Characterizing Risk and Uncertainty and Delineating Appropriate Risk Management Options, with Special Emphasis on Terrorism Risk," *Risk Analysis* 29, no. 4 (2009): 587–600, https://doi.org/10.1111/j.1539-6924.2008.01175.x; and Amos Tversky and Daniel Kahneman, "Judgement and Uncertainty: Heuristics and Biases," *Science* 185, no. 4 (1974): 1124–31.

8. For additional information, see "Collection: Land Transport Security," UK Department for Transport (website), June 30, 2018, https://www.gov.uk/government/collections/land-transport-security; Brian M. Jenkins and Bruce R. Butterworth, *How Sophisticated Are Terrorist Attacks on Passenger Rail Transportation* (San Jose, CA: Mineta Transportation Institute, 2020); Brian M. Jenkins and Bruce R. Butterworth, *Train Wrecks and Train Attacks: An Analysis of Attempts by Terrorists and Other Extremists to Derail Trains or Disrupt Rail Transportation* (San Jose, CA: Mineta Transportation Institute, 2018); and Jeremy M. Wilson et al., *Securing America's Passenger-Rail Systems* (Santa Monica, CA: RAND Corporation, 2007).

9. For additional information, see Jenkins and Butterworth, *Terrorist Attacks on Passenger Rail*.

to extortion or the acquisition of sensitive data.[10] In this respect, the value of chapters 3–5 and 14—each of which addresses various cyber threats and actors—is that they provide the necessary insight to inform a general risk assessment incorporating recognized good practice, whatever the intent of the malicious actor.

Target of Choice or Opportunity?

Potential attackers represent a broad constituency, including ideologues, the disaffected, the disgruntled, and the disturbed. Collectively, they appear drawn to rail locations which exhibit one or more of the following characteristics: (1) easy to access with a limited chance of compromise during hostile reconnaissance or immediately prior to an attack; (2) difficult to secure against the chosen MoA; and (3) likely to result in disproportionate impact and media coverage for a relatively modest attack. Highly motivated and resource-rich terrorists may execute near simultaneous attacks on crowded trains in capital cities, such as the 1995 Tokyo attack, the 2004 Madrid bombings, and the 2005 London bombings (exemplars 11, 10, and 9, respectively). While at the opposite end of the spectrum, the deluded and feeble—invariably better armed than their unsuspecting victims—may simply select the nearest station to their home or work to attack (see exemplars 4–7). Note that in one such case, a lone actor who launched a murderous SBB attack against civilians was eventually held at bay by a man who armed himself with a narwhal tusk and another with a fire extinguisher until armed police arrived and neutralized the attacker.[11]

The breadth of this continuum of attraction represents a challenge to those trying to build a coherent rationale with which to prioritize risk and deploy finite resources. At its broad center lies uncertainty because railways are expansive targets, open to a range of MoA. Particularly during the prodromal phase of an impending attack, a threat actor's motives and intentions may be obscure. As a working assumption, the effort terrorists must exert to achieve a particular objective—and whether they perceive the investment as worth this effort—is likely to influence target selection.[12] This cost-benefit analysis may be especially applicable where the cost to the threat actor is calculated broadly, incorporating the perceived effort required

10. Dimitra Liveri, Marianthi Theocharidou, and Rossen Naydenov, *Railway Cybersecurity* (Athens: European Union Agency for Cybersecurity, 2020), 14–15.

11. "Narwhal Tusk Hero a Year on from London Bridge Attack," *BBC News* (website), November 23, 2020, https://www.bbc.co.uk/news/av/uk-55022920.

12. See Ignacio Sànchez-Cuenca "The Dynamics of Nationalist Terrorism: ETA and the IRA," *Terrorism and Political Violence* 19, no. 3 (2007): 289–306, https://doi.org/10.1080/09546550701246981.

to mount a complex attack, the prospect of compromise, and the chance of achieving the outcome desired. This calculus is perhaps one reason—possibly the reason—why railways are exploited so frequently as vectors for indiscriminate mass casualty attacks, rather than truly strategic operations directed against elements of critical national infrastructure.

From a historical perspective, bombing crowded trains and stations is a well-established tactic of terror. In 1800s London, for example, the Irish Republican Army (IRA) were particularly active on-and-about the railway.[13] The IRA's "S-Plan" ("S" for sabotage) of the early twentieth century noted that transport infrastructure was "probably the most important of all" targets to attack because the resultant "serious dislocation would have a paralyzing effect on every branch of industrial and commercial life."[14] Particularly apparent within the contemporary ideology of jihadis, but also attractive to actors driven by other motives, soft rail targets represent solid targets of choice (see exemplars 9–11).[15] Crowded public places environments where a threat actor's aggressive spirit can overcome a lack of planning, limited technical skills, and meager resources (see exemplars 4–7). There are also numerous examples of weak attacks that failed to achieve the anticipated aim but which, nevertheless, resulted in significant social dislocation and disruption and generated much alarmist publicity (see exemplars 8 and 12). The impact of unrelated low-level incidents can readily combine to elevate anxiety among all rail users and responders alike. In the absence of a resilient approach to risk management, risk aversion in the face of an ill-defined threat can paralyze services, shut stations, place trains out of position (thus impeding the resumption of operations), cause cross-modal disruption, and inconvenience large sections of society (see exemplar 13). A foreseeable consequence of such heightened anxiety is the transfer of risk, particularly if rail users switch to less safe modes of transport.[16]

Risk aversion is typified by a response that: (1) is incoherent, such as evacuating a station after a bomb threat without knowing if a bomb or other hazard exists or where it might be located; (2) exhibits elements

13. Brian M. Jenkins, *The Fenian Problem: Insurgency and Terrorism in a Liberal State, 1858–1874* (Liverpool, UK: Liverpool University Press, 2009), 166–67; and Vivien D. Majendie and Arthur Ford, *Circumstances Attending an Explosion at the Victoria Railway Station, Pimlico, on the 26th February 1884. Report to the Right Honourable The Secretary of State for the Home Department* (London: C. – 3972, 1884).

14. Irish Republican Army, *"S"Plan* (London: Public Records Office, 1938).

15. Wilson et al., *Securing America's Passenger-Rail Systems*, 7–11.

16. For additional information, see Garrick Blalock, Vrinda Kadiyali, and Daniel H. Simon, "Driving Fatalities after 9/11: A Hidden Cost of Terrorism," *Applied Economics* 41, no. 14 (2009): 1717–29, http://dx.doi.org/10.1080/00036840601069757.

of palpable decision inertia, when a generic plan is favored over the management of risk in context (for example, specifying the same evacuation point for all hazards, from fires to bombs and everything in between); and (3) embodies a needless overreaction, one justified against a highly selective and speculative principle of precaution.[17] Consider the example of Umar Farouk Abdulmutallab, a terrorist who sewed explosive material into his shorts to defeat airport screening measures. See the in-depth case study in chapter 6. That esoteric MoA, directed against an aircraft in flight, was extrapolated by some as relevant to mass transit rail. Absent from their analysis was any acknowledgment that access to trains and stations did not involve mandatory screening of any passenger's undergarments. Nevertheless, resources were, albeit temporarily, redirected toward the now readily imaginable but implausible use of intimately worn, wearer-activated, low-yield explosive underpants.[18]

In summary, railway networks are easy to access and difficult to defend, and even a modest attack can bring about levels of disruption out of all proportion to the effort expended. There is every reason to believe that threat actors view railway networks in much the same way. The next section expands the debate by characterizing the nature of the operational rail environment and contextualizing some of the core risk management challenges.

Multifaceted Nature of Railways

Complexity

The term railway is recognized universally, but familiarity is not always synonymous with understanding. Complex and rarely amenable to structural change, rail systems operate not only as single and multimodal hubs and lines of communication, they also facilitate a range of diverse functions. These functions include commercial and industrial activities, retail business premises, office and recreational spaces, and social meeting places. The more complex the system, the greater the range of risk management challenges to be overcome. See chapter 13 for overview of developing risk management strategies. Some of the inherent complexities of railways follow:

17. Paul Slovic et al., "Preference for Insuring against Probable Small Losses: Insurance Implications," in *The Perception of Risk*, ed. Paul Slovic (London: Routledge, 2000), 51–72.

18. "'Underwear Bomber' Umar Farouk Abdulmutallab Pleads Guilty," Federal Bureau of Investigation (website), October 12, 2011, https://archives.fbi.gov/archives/detroit/press-releases/2011/underwear-bomber-umar-farouk-abdulmutallab-pleads-guilty.

- Public space is difficult to regulate. The security architecture of railways means many established technological fixes used elsewhere are: (1) not feasible because the physical structure is difficult to modify; (2) not practical because the established culture governing rail usage is entrenched; or (3) not even desirable because of foreseeable adverse consequences of crowding (from slips, trips, and falls to the inadvertent generation of an additional killing zone). The analysis of aviation security in chapter 6 contains many interesting parallels on this point.
- The identity of those who use the system is mainly unknown. From a practical perspective, it is inconceivable that, within the existing culture of rail travel, passengers would or could arrive hours in advance of a journey and submit to identity checks.
- There are only minimal checks to restrict what people carry about their person and in their luggage. Due to the huge throughput of passengers and foreseeable risks of impeding free movement, such as crowding in confined spaces, restrictions are difficult to enforce.
- Managing high volumes of timetabled rail traffic (passenger and freight) requires precise and centralized coordination. Especially in terms of acts that are malicious, this precision can be exploited to create single points of failure. Railway systems cover a huge geographic footprint, and boundaries between the railway and the surrounding environment—and even between different elements of the same network—can be difficult to define. Without care, it is possible to exacerbate working relationships and create unhelpful turf wars and jurisdictional issues.
- Much of the dispersed infrastructure is accessible, with minimal physical impediment, via authorized and unauthorized points of access. The LoR is often secluded and problematic to secure against even casual trespassers. In terms of passenger movements, network access via a low-risk station will almost certainly facilitate the unhindered movement to higher-risk locations.

- There are multiple normative and derivative stakeholders, and no single hierarchy of priorities is likely to exist. This dynamic is particularly noticeable when rail services traverse local, regional, or national borders.
- Passengers do not represent a cohesive group. The range of behaviors deemed normal can be very broad and the individual or group response to system instability can be unpredictable. As core stakeholders, they are notoriously difficult to engage, and, without great care, risk reassurance activities could provoke risk arousal instead.
- Confirmed or suspected terrorist incidents are invariably net consumers of CT resources. The response outlined in exemplar 9, for example, required the prolonged redeployment of assets drawn not just from London but from across England, Scotland, and Wales.

Regulation and Political Direction

Railways operate within a framework of political direction at the local, regional, and national levels, and face other competing commercial imperatives as well. While unlikely to impact risk management activity as part of the immediate post-attack response, political edicts can constrain operational decision making in circumstances when an evolving and broader policy objective expedites or delays the return to a business as usual (BaU) posture. Such contingencies—where political, operational, and commercial imperatives merge—benefit from being planned in advance and agreed in principle by all stakeholders. That is to say, it is desirable to achieve a broad consensus before foreseeable events come to fruition and certainly before any propensity toward risk aversion stifles what opportunities do exist. Examples of this approach include the use of decision-making templates or memoranda of understanding setting out the conditions needed to close a rail operation in response to a direct malicious threat or an adjacent incident, recommence operation if that threat then fails to materialize, and manage the return to normality after a threat has been neutralized.

Policing and Security

With respect to those practices known as policing and security, when roles and responsibilities become confused, a vulnerability exists. Policing, in its purest form, means maintaining the peace and upholding the rule of law. Active policing is certainly capable of countering terrorism,

and of deterring, detecting, or displacing its adherents, but not in isolation.[19] Success is more likely where the functional structure is optimized toward policing the railway as a specific task, and where organizational memory is accumulated and retained. It is less apparent when the railway is simply another constituency within the general policing effort. In contrast, security is a broader concept that includes the range of physical and procedural measures necessary to make an attack less likely and any response more efficient. As a BaU activity, security is likely to be industry-led—with or without the direction of a regulator or police—but always coordinated within the overall CT effort. Benefit, in terms of public safety and managing societal risk, is maximized when these endeavors operate seamlessly. This desirable outcome is most likely to occur when information necessary to manage identified vulnerabilities is shared through formal channels, involving conduits for both top-down and bottom-up communication. By necessity, sharing sensitive information is likely feasible only within a need-to-know framework involving trusted partners. In this respect, the imposition of artificial barriers—such as security groups that exclude railway experts simply because they are not security professionals—is predictably counterproductive. See chapter 11 for insights on effective information and intelligence sharing.

Railway networks were conceived at a time when terrorism, though not unknown, was certainly of a lower profile and less potent than it is in the twenty-first century. Many stations are not amenable to prescriptive security regimes, such as those involving physical searches and a high reliance on detection technologies. The use of fixed-point detectors, as a particularly problematic example, is predicated upon the existence of a coherent operational requirement that addresses questions such as: How will passenger throughput be maintained while detection data are being processed or exploited? If signal processing is not instantaneous, where might the subject of a positive result be when the alarm is noted, and how will the target individual then be identified and located without delay? What is the anticipated outcome if a person-borne IED (PBIED), firearms, or SBB threat is detected, but the threat actor has already reached a target-rich environment? If the confidence level of a detector technology is known to be below 100 percent, then what value does the system have? What arrangements, such as additional staffing and the associated training burden, physical security, and blast mitigation, are necessary to deal with positive results? There are also the broader

19. *Transport Security: Travelling without Fear: Oral and Written Evidence before the House of Commons Transport Committee*, HC 191 (2008) (memorandum from the British Transport Police), 89–92, http://www.publications.parliament.uk/pa/cm200708/cmselect/cmtran/191/191.pdf.

questions of: who sets the standards for use? Who enforces the standards? And, not least of all, who pays capital and running costs?

Media Impact

Attacks directed against mass transit rail, because of the nature of the hazards they expose and the "dread risk" they represent, are readily comprehensible to stakeholders and the wider population.[20] Such events—because they are imbued with the salience of mortality—invariably generate mainstream and social media coverage, which represents an additional attraction to threat actors.[21] Even in 2005, viewing figures released by the BBC indicated that more than 50 percent of the British population watched *BBC News* coverage on the day of the 7/7 London bombings (see exemplar 9).[22] The observation is of particular relevance now because all-pervasive camera phone footage and instant messaging have become mainstream sources of information. The physical attack focuses attention where rhetoric alone may fail to gain traction, and the newsworthy incident allows threat actors to display power in a manner that is visceral and unambiguous.[23] Moreover, public interest and outrage (see terms in figure 7-1) drive additional exposure, thus delivering a propaganda bonus. Each new attack serves as a reminder of previous atrocities, and all are linked by the common themes of the railway and the vulnerability of its passengers. This ready association can be manipulated further through the tactic of binary terrorism: the combining of physical attacks with preceding or subsequent threatening communications conveying false information.

Terrorist signaling of this type reinforces the spectral nature of the threat to mass transit rail (see exemplars 3 and 13), often becoming a vector by which public attention is diverted from the atrocity toward perceived failures in the response. Recognizing that such exposure is a core element within the threat actor's motivation further reinforces the importance of avoiding overreaction, minimizing disruption, and expediting the return to normality by reinstating rail services as quickly as is feasible. This outcome is achieved most effectively

20. Slovic et al., "Preference for Insuring," 137–46.

21. Aaron M. Hoffman et al., "How Does the Business of News Influence Terrorism Coverage? Evidence from *The Washington Post* and *USA Today*," *Terrorism and Political Violence* 22, no. 4 (2010): 559–80, https://doi.org/10.1080/09546553.2010.493778; and Daphna Canetti-Nisim, et al., "A New Stress-Based Model of Political Extremism: Personal Exposure to Terrorism, Psychological Distress, and Exclusionist Political Attitudes," *Journal of Conflict Resolution* 53, no. 3 (2009): 366, https://doi.org/10.1177/0022002709333296.

22. *Transport Security: Travelling without Fear*, 153.

23. Sophie A. Whiting, "'The Discourse of Defence': 'Dissident' Irish Republican Newspapers and the 'Propaganda War,'" *Terrorism and Political Violence* 24, no. 3 (2012): 483–503, https://doi.org/10.1080/09546553.2011.637587.

when risk is assessed in context, when the process of assessment is inclusive (and not needlessly or cynically selective), and when uncertainty is recognized as an inevitable consequence of dealing with a malicious act.

Plausible Methods of Attack (MoA) in the Rail Environment

This section develops the themes exposed thus far by situating vulnerability in context by using incident exemplars. These exemplars illustrate MoA that are plausible and representative, rather than just recent. Collectively, the exemplars provide support for the contention that countermeasures based upon faulty specification of the problem may be inefficient, subject to a range of unintended adverse consequences, and problematic to defend when subject to legal challenge. In contrast to what follows, this section commences by considering an event unrelated to terrorism in every aspect except in the minds of some of those present and the media sources to whom they spoke.

Fear of Terrorism

Exemplar 1: Exploding E-cigarette on the London Underground (2014)

At approximately 9:00 a.m., a male passenger on a tube train at Chancery Lane was surprised when his man-bag produced a loud "pop" and began to issue smoke. Upon looking inside, he saw his e-cigarette, charging via the USB port on his computer, had failed. The battery case had ruptured due to internal pressure and caused the noise and smoke. While he was reassured, to some extent, by what he saw, fellow passengers were not. The following is taken from an effervescent article in the Huffington Post:

> Passengers screamed and ran out of Chancery Lane station in central London while others were in tears following the security scare at 9:30am. . . . "You just assume the worst, don't you," [an eyewitness] said. "I didn't know to be honest what had happened."[24]

Platform staff were alerted and talked to the surprised owner of the bag. He was behaving normally and offered a logical explanation for events. He showed staff the charred remnants of his e-cigarette and his computer.

24. Jack Sommers, "Suicide Bomb Scare at Chancery Lane Tube Station as Passenger's Laptop Overheats and Gives Off Smoke," Huffington Post UK (website), June 19, 2014, http://www.huffingtonpost.co.uk/2014/06/19/bomb-scare-chancery-lane_n_5510544.html.

There was no fire and no smoke visible by this time. There were no indications of a deliberate explosive event or that the e-cigarette had been modified or adapted to create a weapon. Frontline staff with whom he spoke and who checked the CCTV footage were satisfied this was a minor accident and, in accordance with their security awareness training, they did not report the event as suspicious behavior. The train service recommenced with minimumal delay.

Simultaneously to the pragmatic process of assessment undertaken by staff at platform level, passengers who had self-evacuated were tweeting their experience of being "blown-up on the tube." These reports led to media interest directed at police and the rail operator. Due to the nature of the initial information, and what was appearing on social media, the incident also attracted the interest of specialist CT units. One witness reportedly said she heard a massive bang and people shouted "bomb" as smoke started coming from the backpack.[25] Yet, as was known by staff and police at the station, there was neither a backpack nor a massive bang. Indeed, there was no bomb.

The early stages of a railway accident—particularly one involving fire, explosion, or derailment—may be indistinguishable immediately from an act of terror. The ability to recognize key characteristics from early reports from a scene is a concrete reason why a close working relationship between rail operators and CT responders is desirable and beneficial. Exploiting specialist knowledge concerning how the system should operate—and therefore what is undesirable but normal, as opposed to the genuinely unexpected and suspicious—is invaluable in terms of the collective ability to disaggregate signal from noise (see the comments regarding early reports of an explosion in exemplar 9). Disruption to services will invariably attract media attention and the fear of terrorism may induce a form of contagion among dispersed rail users, provoking well-intentioned but false reporting over a wide area. Developing a security culture, in which frontline staff are invested, is one means of managing this contingency, as exemplar 8 will illustrate.

Sabotage and Attacks against the Line of Route (LoR)

Most occurrences of mechanical or electrical sabotage of rail infrastructure tend to be criminal in nature, rarely driven by a terrorism-inspired methodology. Given the vulnerability of power supplies, signaling equipment, points, and the permanent way generally, this may be considered surprising. There are examples of single-issue groups and lone actors putting items on the track

25. Sommers, "Suicide Bomb Scare."

or attempting to short or burn electrical cables.[26] In 1995, an entity in the United States styling itself "Sons of the Gestapo" tampered with running lines at a remote location. The saboteurs rigged electrical connections to indicate, erroneously, that the track was intact. As a result, a locomotive and 12 cars derailed, and three carriages propelled 10 meters into a ravine. One person was killed and 60 injured. The human tragedy notwithstanding, this event is cited most often because of its rarity rather than its impact. More recently, jihadi publications promoted a homemade derailer: a device fabricated from concrete and steel and weighing only a few kilograms. This call to action appears to have failed as no trains were derailed (see exemplar 2). From a practical perspective, and in contrast to the placement of IEDs (a potentially rapid and more effective attack option), mechanical sabotage may be time consuming, require the participation of several perpetrators on-site, and, possibly, some insider knowledge.[27] It is perhaps for this reason that explosives remain the weapon of choice (see exemplar 3).

Exemplar 2: Specter of the Jihadi Derailer

In 2017, a jihadist publication highlighted the desirability of derailing trains in the United States and Canada using a purpose-made derailing tool placed along the LoR.[28] The tactical use of such a device, however—including optimal placement, attack timing, and concealment considerations—was not addressed in detail. The method also required notable effort in terms of the object's construction, involving the casting of a concrete wedge reinforced with preformed steel. Despite initial excitement within the media and some intelligence circles, it became apparent that simpler and more plausible derailment options existed, as the rail accident data demonstrates. It was also far from clear whether the object as described was fit for the purpose.

Exemplar 3: British Experience of LoR Attacks

From 1991–2001, the IRA was responsible for almost 70 railway-related incidents, 30 percent of which were directed toward LoR targets. Aside from the running lines, attacks also involved signaling equipment, associated items of infrastructure, and, less frequently, and unsuccessfully, bridges. In most cases, bomb placers used the isolation of the locations selected to their tactical advantage. Concealment of devices, if any, was rudimentary. In one case, an IED that was simply laid on top of a ballast, but in close proximity

26. Jenkins and Butterworth, *Train Wrecks and Train Attacks*, 1–2.

27. Brian M. Jenkins and Bruce R. Butterworth, *Long-term Trends in Attacks on Public Surface Transportation in Europe and North America* (San Jose, CA: Mineta Transportation Institute, 2016), 12–13.

28. Jenkins and Butterworth, *Train Wrecks and Train Attacks*, 3–5.

to a rail, detonated under a train that had just been brought to a temporary halt. It is also worth noting that because of the IRA's use of time-delay devices and the possibility of service variations within timetabled movements, exactly where trains would be at the time of detonation could not be known. While derailment was not the specific intention of the bombers, it is pertinent to note that on several occasions the possibility of a high- or low-speed derailment seems to have been a matter of chance.

While some high-profile and highly trafficked locations were chosen, many other targets were remote and of marginal significance. What was apparent is that the bombers made use of the strategic roads network to travel to rail targets, often gaining access where an existing trespass point and off-road parking coincided. This type of analysis was generated through local railway knowledge and post-incident investigation. It proved particularly useful when responding to ambiguous bomb threats, allowing the prioritization of areas of interest and providing a focus for response activity. Conceivably, it was the "hardening" of the station environment—including the deterrent and evidential impact of CCTV, station and train checking regimes, and proactive police patrols—that displaced the bombers and redirected their attention. In the British experience, such measures worked well in prioritized public spaces. With respect to the vast number of remote and secluded locations along the LoR, such measures were less practical.

In other parts of the world, it is apparent that those intent on mass murder often used command-initiated devices and IEDs triggered by train movements. A command element—such as a wire, radio or phone signal, or some form of sensor—ensured a moving train was at or adjacent to the attack locus and approaching at high speed. Chosen locations have often exploited the momentum of the train with carriages being propelled into a more hazardous situation. Notable in countries engaged in, or proximal to, civil or local conflict on a large scale, is armed assault against any survivors. Variations on this theme occurred in Thailand, the Indian subcontinent, Egypt, Algeria, and South Africa.[29]

Physical Assaults against People

As a tactic of terrorism, assaulting individuals incrementally and with lethal force is a relatively recent railway-related MoA. It may have been inspired by beheading propaganda of the type available online, but it is reasonable

29. See Jenkins and Butterworth, *Train Wrecks and Train Attacks*; and Brian M. Jenkins, *Protecting Public Surface Transportation against Terrorism and Serious Crime: Continuing Research on Best Security Practices* (San Jose, CA: Mineta Transportation Institute, 1997), 206–48.

to surmise that attraction lies in the ready availability of large knives, axes, and other such weapons. In China, the tactic was used in a rampage-style attack in 2014, and to some notable effect.[30] In Western Europe and the United States, however, it is clear that just as "rock beats scissors," a police officer's gun beats a terrorist's knife (it is notable that SBB weapons, however, have been used to ambush police officers and steal their firearms). It is also clear that attacks in confined spaces offer some tactical advantage to the threat actor. One notable development, a crude but sometimes effective attempt to reduce the efficacy of any armed police or military response, has been the overt wearing of objects intended to represent a PBIED. The hoax PBIED has proved to be of limited protective value against trained shots, especially when following suitably robust rules of engagement.

Exemplar 4: UK Incident (2018)

During the midevening of New Year's Eve, a lone attacker launched an unprovoked and spontaneous stabbing and slashing attack against two people at Manchester Victoria station. Prior to the assault, the attacker recited Islamic verses. A single police officer was quickly at the scene, but Taser, Captor, and physical restraint techniques proved disappointingly ineffective. Only with support from additional officers was the attacker subdued and arrested. The attacker had two knives, one in his hand and one about his person. During the incident, he repeatedly stabbed the police officer. Even when under restraint, he was noncompliant and maintained a jihadi-style rant. The man reportedly suffered from mental health issues.

Exemplar 5: French Incident (2017)

During the early afternoon of a Sunday in October, a man armed with what was described as two butcher's knives, one of which was concealed up his sleeve, attacked and murdered two young women in the vicinity of the railway station in Marseille. The attacker was dispatched promptly by soldiers on dedicated station patrol duties, supported by a number of covert police assets. The incident was later claimed by the Islamic State (Da'esh). The assailant, a drug addicted low-level criminal, is not thought to have been radicalized within a formal cell structure.

30. Priya Joshi, "Kumming Station Knife Attack," *International Business Times* (website), March 1, 2014, http://www.ibtimes.co.uk/kunming-station-knife-attack-authorities-claim-mass-stabbing-was-organised-premeditated-warning-1438515.

Exemplars 6 and 7: German Incidents (2016)

At approximately 9:00 p.m. on July 18, a lone young male—an Afghan refugee local to the area—attacked passengers on a suburban rail service in Germany. He was armed with an axe and a knife. During a relatively short attack, three people were seriously injured and at least two others received lesser injuries. The train was not operating at capacity and was not, therefore, a target-rich environment, but it was a confined space. Media sources reported the incident as religiously motivated because of what the youth is alleged to have shouted and his possession of an Islamic State (Da'esh) flag. He was shot dead by police when he disembarked the train and assaulted armed officers.

During the early morning of May 10, a single assailant launched an unprovoked knife attack at a suburban railway station in southern Germany. The attacker was a German citizen in his late twenties. At least four people were injured, one fatally. As noted as part of previous SBB incidents, it is alleged the assailant aligned himself to jihad during the commission of the attack. Using batons, police officers were able to overpower the man without firing a shot. It is not clear why the attacker chose a relatively small suburban station or why he attacked at approximately 5:00 a.m., in advance of the peak travel period. It is reasonable to suggest that target selection was idiosyncratic. The man was not familiar with the area and the timing of the attack may relate to the fact that upon his arrival at 1:38 a.m., the assailant was unable to find a hotel, so he simply returned to his point of arrival. Mental health issues were considered significant with respect to the man's behavior.

Collectively, the exemplars illustrate that target selection can be idiosyncratic in nature. Notably, many assailants go unchallenged prior to the commission of the attack, with any odd activity being absorbed within the wide range of behaviors deemed normal for a railway. On the continuum of accessibility, assailants were sometimes drawn to targets that may, as a part of any formal process of threat assessment, be deemed marginal. In each of the cases discussed here, the inherent vulnerability of mass transit rail was redressed by prompt and decisive action by police who were either patrolling the locale or dispatched as a mobile resource.

Improvised Explosive Devices (IEDs)

There are numerous examples of IED attacks against rail passengers, rolling stock, and infrastructure. Where human capital is limited or otherwise valuable, hand-placed or sometimes vehicle-borne, timer-controlled IEDs

remain the preferred MoA. Where the intended target is perceived as difficult to access or otherwise protected, stand-off weapons or delivered IEDs have been utilized. When the bombers intend to give their own lives, whatever the motive, PBIEDs may be encountered. Command-initiated devices have been used where the target was mobile. Variations on the themes identified have included those listed below.

- No-notice attacks (presumably, the absence of a warning is to maximize casualties)
- Attacks preceded by an accurate warning (presumably to minimize casualties but also as a means of claiming a "moral" position should the attack not go to plan)
- Attacks preceded by a warning containing deliberately false information to maximize disruption and/or create a killing zone at predictable evacuation routes and assembly points and divert blame (see socially engineered attacks in exemplar 13)
- Attacks involving multiple devices set to function simultaneously or to a pattern of delay
- Attacks involving a mixture of IED types (for example, high explosive interspersed with incendiary, vehicle-borne IEDs and hand-placed IEDs, and time-delay IEDs protected by anti-handling circuits)

A feature of attacks in stations is that hand-placed IEDs deposited in public spaces were often noticed (sometimes stolen) before they functioned. This reality seems to have led attackers, at times, to deliberately place the devices where they were less likely to be seen and reported quickly (such as in litter bins and lavatory cubicles). Throughout the 1990s in the United Kingdom, instances of bombs being abandoned in public spaces in stations declined sharply as a result of a formalized station checking regime, implemented by trained rail staff working to a directed plan, and the removal of areas of concealment, especially litter bins. This initiative brought an unexpected security bonus in terms of deterrence: litter-picking activities undertaken by staff in high-visibility apparel reinforced the impression of public spaces being "owned." Associated measures involved a structured campaign of risk communication directed at passengers, prioritized police patrols timed to support staff security activities, and enhanced CCTV coverage.

Exemplar 8: Low-level/Low-sophistication IED, London (2016)

During the midmorning of Thursday, October 20, passengers traveling on a Jubilee Line train noticed a small rucksack that had been left in clear public view adjacent to the train's double doors. The bag was ignored for several station stops but, at Canary Wharf, a passenger picked up the bag and handed it to the train's driver, who placed it in the cab. Only when the train was moving did the driver become suspicious and, eventually, call police. Following an assessment by specialist officers, an explosive ordnance disposal (EOD) team disarmed a crude but viable timer-controlled IED incorporating a confined low-explosive composition. In summary, this incident includes the following highlights:

- The construction of the IED was rudimentary but did include additional fragmentation.
- The timing of the attack appears not to have exploited the peak travel period and the train was not operating at capacity.
- The bag was left in clear public view to the extent that other passengers watched the bag's owner exit the train.
- In contrast to more successful attacks (see exemplars 9–10), only a single device of low mass was used.
- The item was not subject to any deliberate process of assessment until examined by police.

It remains unclear why the bag was moved, handed to the driver, and then—against training, documented protocols, and expectations—placed within the cab. This action was taken despite the weight being unevenly distributed relative to its bulk and, of greater concern, the visible presence of wiring, a modified electromechanical wall clock, a dry-cell battery, and wires entering a thermos-type flask. Potentially, this was an example of what Feudenburg defined as the atrophy of vigilance, since it occurred during a period in which train bombings in London had become rare events.[31] Utilizing data from London's extensive CCTV network, the arrest of an autistic young man, whose interest in weapons was identified as misguided rather than terroristic, followed quickly. There was, however, evidence presented that the perpetrator had read an online publication of dubious technical merit but, nonetheless, attributed to al-Qaeda.

31. William R. Freudenburg, "Nothing Recedes Like Success—Risk Analysis and the Organizational Amplification of Risks," *Risk* 3, no. 1 (1992): 19, https://scholars.unh.edu/cgi/viewcontent.cgi?article=1071&context=risk.

Exemplar 9: Expansive Attack, London (2005)

On the morning of July 7, four young men traveled from their home counties to London by train. Upon arrival at the rail hub (King's Cross), three of them then boarded different underground trains. Eight minutes into their journeys, and only when the carriages were traveling through the closely confined environment of a relatively shallow tunnel, the first two PBIEDs were detonated. The third explosion happened 500 meters into a very deep section of tunnel. It is estimated that the three explosions occurred over the course of less than one minute. The fourth bomber had been unable to board a train before his accomplices initiated their devices; he had to buy a new battery for the IED he carried, but may also have been a less committed individual. Since train services were suspended, he boarded a bus and detonated his PBIED shortly after the journey commenced. The four explosions occurred over the course of one hour, with the train bombs clustered temporally at approximately 8:50 a.m. and the bus bomb functioned at 9:47 a.m. First reports were initially attributed to an accidental power outage—a recurring and anticipated electrical supply problem—but police officers in close proximity recognized the characteristics of an explosion: the distinctive noise, ground shock, and dust issuing from the tunnels. It was not immediately clear how many trains were affected because multiple reports pertaining to the same incidents were being received. Within 25 minutes, all trains were brought to a halt at platforms, and a full evacuation—in accordance with a well-practiced plan—commenced shortly afterwards. By that stage, camera phone images of the carnage were appearing online.

People displaced from the rapidly closing underground stations formed large crowds in the streets. Some attempted to continue their journeys by taxi or bus while others waited for train services to resume. For some time, it was not clear whether other attacks were pending, and numerous reports of suspicious behavior were received. This uncertainty was heightened when the delayed fourth IED detonated. Due to the rapid evacuation from tube trains, numerous unattended bags required police assessment, further stretching resources and causing additional disruption. The vast majority of unattended bags were dealt with quickly using a railway-specific assessment protocol called H-O-T.[32] This simple heuristic devised in the early 1990s aims to disaggregate noise—the discovery of an unattended bag in an environment where such discoveries are common—from the signal, the possibility that a bag may contain an IED. Using the H-O-T heuristic, trained rail staff or police focus

32. *Transport Security: Travelling without Fear* (2008), 84–86.

on three aspects of known relevance.[33] Is the item hidden? Very few passengers who forget their belongings hide them first. Is the item obviously suspicious in appearance or in the circumstances of its discovery (see exemplar 8)? Is the item typical of what is usually discovered in the environment in question?

Following police search activity and the rapid redeployment of additional officers, elements of the underground network began a phased reopening later that afternoon. In total, 52 people had been murdered and many more injured grievously.[34] As the underground incident scenes were particularly distressing and challenging, those elements of the network affected directly remained closed for several weeks. This disruption was not primarily because of bomb damage, but because of the protocols used to manage the crime scene.

Exemplar 10: Expansive Attack, Madrid (2004)

On March 11, terrorists deployed 13 IEDs on suburban commuter trains converging on Madrid. The bombs, visually indistinguishable from legitimate luggage, were placed as the trains waited at relatively remote suburban stations. Between 7:35–7:40 a.m. (note: timings vary depending upon the source consulted), 10 explosive devices detonated in four trains along the C-2 commuter train line running from Guadalajara to Atocha Station. Each IED was contained within hand-carried luggage. In addition to the 10 bombs that detonated, EOD rendered safe the remaining three devices, including one discovered among abandoned luggage some hours later. The bombs were timer-activated and set to explode after a delay of 30 minutes. With approximately 700 people on each of the trains, the carriages were crowded. Witnesses mention seeing the bags being placed on the train, but none considered the action sufficiently suspicious to report it at the time.

The first explosion (of three on the same train) occurred on a commuter service that had already arrived at Atocha station, killing 29 people and wounding 176 others. A fourth IED was later located on the train and made safe by EOD. Further explosions occurred on another commuter train that was running two minutes behind schedule. This train was moving into Atocha when four bombs detonated, stopping the train approximately 500 meters away from the first bomb scene, killing 59 people and wounding 200 more.

33. Centre for the Protection of National Infrastructure (CPNI), *Recognising Terrorist Threats* (London: CPNI, March 9, 2020), 6–7, https://www.cpni.gov.uk/system/files/documents/b8/40/CPNI%20-%20Threat%20Recognition%20Guide%20-%20WEBv2.pdf.

34. *Intelligence and Security Committee, Report into the London Terrorist Attacks on 7 July 2005* (London: Her Majesty's Stationery Office, 2006), 2, https://www.gov.uk/government/publications/report-into-the-london-terrorist-attacks-on-7-july-2005.

Almost simultaneously, two explosions occurred on a third train, and, sometime later, two further IEDs were made safe by EOD. As a result of these explosions, 67 people died and there were 200 casualties. A little further away from Atocha, a single bomb exploded in a train carriage, killing 20 people and wounding 50 others. During the weeks after the attack, a further 26 people died of their injuries. In total, the bombings claimed over 200 lives.[35]

The inherent vulnerability associated with mass transit rail was clearly relevant in the London and Madrid examples. Bombers entered the rail network at points where security measures were marginal and used the trains to bring the IEDs to more critical locations. The fact that passengers in Madrid watched the bags being abandoned in the carriages in which they were traveling, yet did nothing, indicates a major defect in the prevailing culture of security (see exemplar 8). In contrast to London, train services in and out of Atocha were not suspended, and the crime scenes were dealt with exceptionally quickly. In Spain, a seemingly robust approach to risk management meant minimizing disruption and expediting the return to normality were established political and operational priorities. As in exemplar 9, there was an enhanced security and police presence after the attacks, but, unlike London, military assets were integral to the high-profile policing function.

Quick-acting Noxious Hazard

Quick-acting noxious hazards are associated most closely, but not exclusively, with chemical weapons. To date, terrorists' relatively ineffective and infrequent use of this MoA is difficult to reconcile, especially given both their declared intention and demonstrated willingness to perpetrate mass casualty attacks by other means. When compared to the familiarity with and proven capability of explosives, however, a rationale based upon pragmatism emerges. The seminal example of an incident affecting rail mass transit occurred in Tokyo and is outlined below.

Exemplar 11: Tokyo Metro (1995)

The attack on March 20 was a resource-heavy criminal enterprise involving the production of an approximation of sarin, a volatile nerve agent from the World War II era. When the attackers manually disseminated multiple containers on several trains—traveling on three different lines yet each passing through the same hub station—it created a major incident

35. Glen Segell, "Intelligence Methodologies Applicable to the Madrid Train Bombings," *International Journal of Intelligence and Counter-Intelligence* 18, no. 2 (2005): 221–38.

of massive proportions.[36] The attack commenced at approximately 8:00 a.m. on a weekday morning. Initial reports of people in distress on crowded trains were not immediately recognized as a precursor for what was to follow; the initial belief of the railway control was that the cause must be accidental. Despite understandable uncertainty within the response community, it is noteworthy that local news reported the incident by 9:00 a.m., and international news had picked up the story by 9:12 a.m. Approximately 90 minutes after the sarin had been released, an attack was declared and enhanced security put in place. Over the course of the day more than 4,000 people sought hospital treatment, some carrying contamination with them into hospitals. Such was the confusion that one contaminated train completed its journey and then began the return trip toward the incident scene.

The relatively low efficacy of the nerve agent, combined with its inefficient release meant anticipated lethality was reduced considerably. The sachets in which it was contained were punctured manually—at which point the perpetrators decamped—and dissemination then relied upon evaporation of the pooling volatile liquid. Contemporary figures vary, but it seems likely that the overwhelming number of self-reporting casualties were, in fact, the "worried well." In contrast to the multiple train bombings in Madrid and London, the death toll was relatively small, as eight people died within the first 24 hours from among 1,000 who displayed clinical symptoms. Over the years, a number of victims have died as a long-term consequence of being poisoned. The incident highlighted the relative unpreparedness of the railway and responders when dealing with a highly volatile fast-acting and unrecognizable chemical hazard—specifically, the challenges of detection, identification, and monitoring. These difficulties meant many responders became contaminated unknowingly and contaminated evacuees were neither identified nor controlled. These challenges resulted in notable cases of secondary contamination among taxi drivers, people at street level, and hospital staff.

Since the Tokyo attack, the chemical MoA has become a mainstream CT concern. Chemical attacks are generally considered in concert with biological, radiological, and, somewhat incongruously, nuclear hazards. Combining these four unique hazards into one acronym, CBRN, masks the very different challenges and relative plausibility associated with the discrete hazards the acronym seeks to encompass. Had the Tokyo attack incorporated

36. Javed Ali et al., *Jane's Chemical-Biological Defense Guidebook* (Alexandria, VA: Jane's Information Group, 1999), 219–25.

a biological hazard with a longer incubation period (measured in days or weeks), the scale of the subsequent public health emergency could have been overwhelming. Apart from a jihadi propaganda effort (the so-called Mubtakar [sic]), similar chemical attacks have not occurred in the United States or Europe (though, somewhat unhelpfully, some databases record incidents involving irritant sprays under the heading of chemical attack).[37] On a smaller scale, nerve agents have been used for political assassinations unrelated to railway travel.

In the Tokyo attack, it is clear that organizational surprise was a significant factor in failing to respond to what was being reported from the numerous incident scenes. Additionally, the nonpersistent nature of the hazard greatly aided the swift return to normality, especially when compared to the extended duration of the decontamination effort following the anthrax (biological) contamination of buildings in the United States in fall 2001.[38] Practical countermeasures adopted to date focus extensively on ensuring potential victims (1) identify the signs and symptoms of attack, (2) recognize the immediate need to move into fresh air and remove any contaminated items of clothing (the existence of challenges relating to modesty, inclement weather, and the possibility of noncompliance notwithstanding), and (3) await specialist advice and do not act as a vector for secondary contamination. In recognizing the need to expedite the return to normality, some underground rail systems routinely use monitoring equipment to profile the environmental conditions at vulnerable locations. This activity provides baseline readings against a criterion of what is normal for that operating environment and increases user familiarity with detection, identification, and monitoring apparatus.

Firearms

Large-scale incidents involving firearms on trains or in stations are rare events. Increasingly, jihadis have incorporated automatic weapons in an offensive role during attacks at airports and some marauding incidents, but not usually at stations. Some threat actors have carried side arms for personal protection, but rarely as the primary weapon of attack.

37. "Mubtakar Improvised Cyanide Gas Device Warning," Department of Homeland Security (website), November 29, 2010, https://publicintelligence.net/ufouo-dhs-mubtakar-improvised-cyanide-gas-device-warning/.

38. See National Research Council, *Reopening Public Facilities after a Biological Attack: A Decision Making Framework* (Washington, DC: National Academies Press, 2005), https://doi.org/10.17226/11324.

Exemplar 12: Thalys Train Attack, Belgium and France (2015)

On August 21, an armed assailant emerged from the lavatory of a train in service on the Amsterdam to Paris route. He had an assault rifle with a folding stock (enabling him to hide it within hand luggage), several spare magazines, a self-loading pistol, a knife, and a container of flammable liquid. Given the very limited options available to them, but showing great courage, passengers on the train attempted to restrain the man. One person suffered knife wounds and another was shot with the pistol, but there were no fatalities. The gunman was unable to prepare his assault rifle to fire reliably because he failed to clear a stoppage. As a consequence, he was overpowered by the passengers he set out to kill. In reviewing this incident, it is apparent that the element of surprise and the selection of unarmed people in a crowded and confined space provides shooters with a clear tactical advantage. This advantage extends over victims and armed responders alike, especially if the exact location of a train in service cannot be determined quickly.

This incident raises serious questions about the validity of official advice suggesting that would-be victims should not challenge an armed assailant but should run and hide instead. How passengers should follow this advice on a train in service is not clear. Similarly, despite creditable improvisation by some members of the train crew, the contingency had not been foreseen and no countermeasures were in place. In particular, some of the train staff seem to have been overwhelmed by the event and played no active part in its resolution.

There may be no single reason for the low uptake in a firearms-based MoA against rail targets, particularly given the exploitation apparent in other environments, but it is possible to speculate that no-notice IEDs are perceived as more effective. The increasingly efficient and focused action of better-armed and better-trained responders may also be regarded as a more credible deterrent than it once was.

Social Engineering

This chapter has already alluded to the issue of social engineering—that is, to induce a false belief thus causing intended victims to work against their own best interests—by highlighting examples when threat actors wore hoax PBIEDs to distract both victims and responders. Encountered more frequently is the use of threatening communications, such as messages suggesting an attack is imminent and people must be evacuated. This tactic is also encountered as a feature of extortion scams and

only very rarely are such messages from terrorists. Even when a causal link is established, the information these messages contain is of variable quality or even deliberately misleading.

Exemplar 13: IRA Binary Terrorism, United Kingdom

The following example, taken from a contemporary news report, is representative of valid bomb threats. It illustrates the range of difficulties, often encountered simultaneously, when police or other risk managers attempt to divine meaning from information.[39] A series of similar, but not identical, communications marked the commencement of a campaign of binary terrorism lasting more than 10 years. During that period, British police dealt with 10,000 railway-related bomb threats yet fewer than 1 percent resulted in the event they claimed to presage. The first message of the new campaign—received after one IED had already detonated—was reported thus:

> Before a bomb exploded in a trash bin at Victoria Station during the morning rush hour, a caller with an Irish accent told authorities, "We are the Irish Republican Army. Bombs to go off at all mainline stations in 45 minutes[.]"... At least 19 false calls were received after the first explosion, which prompted British Transport Police to begin a search of stations . . .[40]

The threatening communications, delivered early on a weekday morning, specified the targets as all mainline stations in London, but only two stations were targeted (and it remains unclear why one detonation occurred before the first threat was received). Encoding the message in this way may have been intended to focus police attention on central London terminals without disclosing which ones—an act possibly intended to maximize disruption while still supporting a claim of good faith. Its meaning, however, was interpreted in accordance with common and accepted railway usage, and the spatial descriptor was taken to mean any station on any mainline. This ambiguity, coupled with the imprecise reference to "in London," substantially increased the number of potential locations affected from single figures to several hundred.

39. Loet Leydesdorff, "The Communication of Meaning and the Structuration of Expectations: Giddens' 'Structuration Theory' and Luhmann's 'Self-Organization,'" *Journal of the American Society for Information Science and Technology* 61, no. 10 (2010): 2138–50.

40. Karin Davies, "One Killed, 40 Wounded in Suspected IRA Attack," *United Press International* (website), February 18, 1991, https://www.upi.com/Archives/1991/02/18/One-killed-40-wounded-in-suspected-IRA-attack/9583666853200/.

Also of significance was the effect of noise on the process of decoding and understanding what the message really meant. A third party received the threat before it was then passed to police—a process that accentuated uncertainty about the fidelity of the message. In a different incident, it was noted that the wording of a bomb threat was passed verbally from the original call-taker to a manager whose recollection of the words used was then transcribed by a third person, and only at that point did the police receive the message. In that example—and as would often seem to be the case after an actual bombing—police were required simultaneously to deal with noise from competing threats of unknown origin, all of which proved to be hoaxes.

To the threat actor, the value of this MoA is that it exploits any predisposition toward risk aversion and, as such, represents a force multiplier. Acute vulnerability is especially notable where contingency plans are inadequately developed or generic and where a principle of precaution is the default position. In numerous examples of bomb threats against rail services, the outcome has been significant disruption, sometimes resulting in casualties attributed exclusively to uncoordinated evacuation activities.[41] This situation can precipitate copycat threats and result in the unregulated transfer of risk. In a smaller but significant number of cases, decision inertia has led to people being moved from a place of relative safety into a place of acute danger. This undesirable outcome occurs when decisionmakers do not seek to differentiate between valid and hoax communications—that is, to estimate the credibility of the threat the being asserted.

Mixed-methods Attacks

In one of the two examples below (London), an attack incorporating mixed weapon types had been anticipated and resources allocated accordingly within a coordinated response plan, especially the use of armed police.[42] The major station involved had measures in place to restrict vehicle access and a tested multiagency contingency plan. Most notably, the responders to the scene were better armed than the marauding terrorists. In the other (Mumbai), the MoA was unexpected and delivered by heavily armed, well-trained assailants with speed and focused aggression. The different outcomes are readily apparent.

41. Freudenburg, "Nothing Recedes like Success," 1–2.

42. Chief Coroner, *London Bridge and Borough Market Terror Attack*, 17–18.

Exemplar 14: Adjacent to London Bridge Station (2017)

During the late evening of June 3, pedestrians in the vicinity of London Bridge station were struck by a transit-type van. Three men exited the vehicle and then moved toward Borough Market, a commercial and residential area adjacent to the station. They immediately commenced a series of physical assaults, using long-bladed weapons that had been secured to their wrists with adhesive tape. No firearms were discharged and no IEDs deployed. The men were seen to be wearing bulky items on their belts—items assumed to be PBIEDs—but which were later found to be hoaxes. All assailants were neutralized by armed police within approximately 10 minutes. During the rampage, seven people were murdered and 48 injured, including an unarmed patrolling officer of the railway police. London Bridge station did not become the focus of the attack, but was quickly closed and evacuated, and given its proximity to the crime scene, the station remained closed until the next week.

Exemplar 15: Central Mumbai Station (2008)

The attack of November 26 involved near simultaneous incidents initiated by highly mobile, highly motivated, well-trained, and well-armed threat actors. Multiple teams attacked several locations over a very short period of time. The MoA included: armed assaults incorporating the use of hand grenades, carjacking, drive-by shootings, timer-controlled IEDs, directed assassination of armed police and selected foreigners, indiscriminate killings, building takeovers, and barricade and hostage taking. Relevant to this chapter are the events at Chhatrapatig Shivaji Terminus (CST), Mumbai's central railway station. This station was a predictably crowded and target-rich environment, and a small number of attackers was active for approximately 25 minutes. Commencing in the late evening, more than 50 people were killed and more than 100 injured, with the majority of casualties taking place in the first six minutes of firing. Casualties at CST represent approximately one-third of all those recorded throughout this incident. This high number of casualties highlights the vulnerability of a crowded station where there are few places to provide immediate cover from view or from kinetic effects. The series of attacks extended across several days, nearly 60 hours in all, before the attackers were neutralized. Despite initial claims that more than 20 terrorists were roaming the capital, there were actually 10 active shooters, nine of whom were killed.[43]

43. CNN Editorial Research, "Mumbai Terror Attacks Fast Facts," *CNN World* (website), September 18, 2013, https://www.cnn.com/2013/09/18/world/asia/mumbai-terror-attacks/index.html.

This attack was expansive and novel when measured against previous incidents in Mumbai, which involved similar large-scale loss of life but through the use of IEDs. The speed and aggression of the attack overwhelmed the initial responders at the railway station; they were not expecting an armed assault, were not trained to deal with the situation they confronted, and were not armed or protected to take on the attackers on equal terms. It is notable that attackers actively avoided those who could respond in kind, breaking contact quickly when encountering serious opposition. CST exhibited the inherent vulnerability already discussed. High levels of uncertainty meant responders could neither immediately comprehend what was happening on the ground nor implement a coherent response to it. The delayed response was inevitable because of the large number of reports the police had to process. During the initial confusion at CST, more people entered the attack locus, both as pedestrians and on inbound trains that could not be stopped. By the time responders were organized, the marauding attackers had moved on to their follow-on targets.

Developing the Lessons Available

The case study section illustrates a variety of established MoA to which railways remain vulnerable, and to which terrorists have gravitated in the past. Whether they will do so again is beyond the scope of this chapter except to note that threat actors often return to MoA perceived to have worked for others, especially where exposed vulnerability remains unaddressed. Variations noted in attack methodology, when they do arise, tend to be in terms of scale, skill, or scope rather than novelty or originality. This situation is problematic, because the vulnerabilities that threat actors exploit are integral to how railway systems were designed to operate. With the notable exception of the Tokyo subway attack, the following themes are discernible:

- Attacks against people have most notably involved explosives, often augmented by the addition of fragmentation. Firearms and other close quarter attacks involving SBB weapons have come to the fore in recent years, and determined attackers seem to mix-and-match MoA as a deliberate tactic.
- Attacks against fixed infrastructure targets involved primarily explosives and sometimes flammable materials. Mechanical sabotage has been relatively uncommon.

- Attacks against trains in service involved explosions in carriages or along the train's route. In some regions of the world, firearms have been used to attack survivors and, occasionally, rake passing trains.

It must be assumed that terrorists will continue to build on their strengths and exploit weaknesses in the governments and institutions of NATO member states and partner nations. Looking beyond attacks within, or proximal to, war zones where terrorists' capability and capacity may be extensive, analysis of rail incidents suggests substantial mass casualty events represent relatively rare occurrences. This is not to suggest that aspirations among would-be or established terrorists are not high or that unexpected and transient successes might occur. Incidents which combine established MoA, when encountered simultaneously or incrementally, remain of particular concern. They exist as a warning against developing risk assessments that only consider one specific MoA (see exemplar 15) without taking into account how to respond when faced with situations involving multiple MOA. Having separate contingency plans for SBB, firearms, and IED attacks is very different from responding to a scenario in which they occur simultaneously. Terrorists consider cunningly elaborate MoA, most of which overpromise, under deliver and, after considered assessment, are often found to be implausible.[44] In this regard, care must be taken not to overestimate the magnitude of the risk management challenge when novelty generates heat rather than light. The efficacy of the countermeasures deployed is affected by a range of variables, and no single option represents a panacea, hence the need for an integrated approach. As is readily apparent, if high levels of security only apply to some central locations, then there is little to prevent attackers from entering a network at its fringes and using trains to access targets of choice (note exemplars 9–10). Realizing value in relation to any investment in physical security relies heavily upon the quality of the concept of operations and the defined operational requirements within which it resides.

In considering the collective substance of the exemplars provided here, there is a strong argument against target hardening that results in open railways becoming fortified citadels, even if such a change were possible without destroying the societal benefits rail users have come to expect and upon which economies depend. Instead, to maintain a credible BaU posture in the face of an uncertain threat, spaces should be owned by those who work in them and use them, including dedicated security and police assets.

44. For additional information, see "Australia Terror Accused Discussed 'Kangaroo Bomb,'" *BBC News* (website), January 28, 2016, http://www.bbc.co.uk/news/world-australia-35425665.

See the recommendation in chapter 6 for airport community security programs. Developing a strong culture of security can constrain a threat actor's freedom of movement and increase the likelihood of recognizing suspicious activities. As a consequence, reports of concern become more meaningful because of the context with which they are explicitly associated (note exemplar 1). Moreover, the organizational response will exhibit greater cohesion and resilience because stakeholders share an approximation of the same worldview and are invested in the risk management process (an observation that also represents a strong argument in favor of specialist rather than generalist railway policing). The scale of this task should not be underestimated. As stressed throughout the chapter, integral to the process of developing a context-specific StRA is the ability to reconcile the purpose of rail operations with a rational appreciation of exposure to plausible threats. This is a task that requires significant input from those who understand how railways operate as public transport networks, not just as abstract security challenges.

Conclusion

Utilizing case study data drawn from the rail environment, this chapter has discussed a number of contemporary debates concerning proportionate, pragmatic, and effective options to manage terrorism-related risk. It has addressed the nature of the inherent vulnerability of rail operations and identified the need to develop a clear understanding of what risks are relevant and in what context. The process outlined requires the inclusion of multiple stakeholders, some of whom may adhere to differing precommitments. This reality reinforces two factors: (1) the requirement for an inclusive, coherent, and transparent process of risk assessment, and (2) the benefit of developing a risk management doctrine—including templates to guide active decision making under conditions of uncertainty—agreed upon in advance by relevant stakeholders. Countering such a diverse threat against such a diverse target involves:

- Adopting a pragmatic and, as far as is possible, transparent and inclusive approach to the strategic assessment of risk.
- Ensuring the resultant StRA is informed by professionals with a clear grasp of how best to achieve the primary purpose of networked rail services.
- Devising a concept of operations that matches the context-specific vulnerabilities and plausible threats identified.

- Developing a security culture within which all stakeholders are able to share an approximation of the same worldview.
- Not allowing uncertainty to ferment decision inertia.

Neither the attractiveness of rail operations to terrorists—especially the target-rich nature of crowded trains and heavily trafficked stations—nor the challenges of managing risk without compromising utility have been underestimated here. All options require application in context, particularly in the tightly coupled environment of a rail network, because any isolated quick win is rarely synonymous with a permanent fix. Within prevailing cultural norms and societal expectations, most railways are structured as "turn-up and go" transport providers, and this functional reality is not amenable to structural change in the short or even medium terms. As one senior figure within the British Cabinet Office noted after the 2005 London bombings, when explaining why terrorist bombers had faced no identity or baggage checks when entering the rail network, "You have to be able to travel if the purpose is to travel."[45] This statement implies that checking the status of every passenger while facilitating almost eight million journeys each day was an impossible task. That statement was a refreshingly frank recognition that there will always be dynamic tension between the needs of the mass transit operator—including the practical constraints they face—and those charged with countering terrorism. This chapter has set out one approach to reconciling such differences as a means of managing risk more effectively and maintaining a BaU posture more credibly.

45. Sir Richard Mottram GCB, Security and Intelligence Coordinator, Cabinet Office, *Transport Security: Travelling without Fear* (2008 statement), 6.

– 8 –

Water Sector Resilience and the Metropolitan Washington Case

Steve Bieber

In most urbanized societies, water is taken for granted and little thought is given to how fragile the supply of this vital resource can be. A water emergency, however, such as a treatment plant outage, a water source contamination event, or natural disaster has the potential for significant disruption to society and the infrastructure that depends on water to function. Most other sectors of critical infrastructure, as well as activities of daily living, are highly dependent on the water sector. As a result, consequences of a water emergency can be significant and may occur immediately without notice depending on the nature of the event. Thus, the security and resilience of the water sector is a key component of a nation's civil preparedness that can have military and international implications as well. Terrorist threats to water delivery or contamination of water sources as a terrorist act can impact a nation's ability to move and sustain its military forces, and project military power when required. From the perspective of the North Atlantic Treaty Organization, threats to the water sector in one member state could have ripple effects that limit or diminish NATO's military mobility and force projection in support of its essential core tasks.

Therefore, it is important to understand water sector risks and find ways to effectively mitigate them. While this chapter focuses on the US water sector and uses a case study from one of its most important metropolitan areas, the chapter provides a helpful framework for other Allies and partners to understand, adapt, and employ to their specific circumstances.

To that end, this chapter includes five main sections to: (1) provide an overview of the water sector; (2) identify risks and threats to the water sector; (3) outline key steps in resilience planning; (4) illustrate challenges and solutions to security and resilience initiatives using a case study from Washington, DC; and (5) offer recommendations for developing water-sector security and resilience.

Understanding the Water Sector

Water for drinking, bathing, electric power generation, manufacturing, and other uses is a precious commodity. The amount of available freshwater on earth is estimated to be only one half of one percent of all water on earth, as figure 8-1 depicts.[1] When it comes to the amount usable for drinking, if the world's supply of freshwater was 100 liters, then the usable supply of freshwater would only be about 0.003 liters or one-half teaspoon.[2] In industrialized countries, the term on tap is often used when talking about things that are as freely and readily available as water. In the United States, water is inexpensive and abundant, and the average per capita demand is about 80–100 gallons per person per day.[3] Most of that water is used for non-potable uses such as cooking, bathing, toilet flushing, and doing laundry.

The US Centers for Disease Control and Prevention identifies the availability of clean, safe water and sanitation as among the top ten public health accomplishments of the past century.[4] Engineers Abel Wolman and Linn Enslow determined the correct formula for treating public water supplies with chlorine in 1919, and approximately 10 years later, the widespread adoption of chlorination, filtration, and improved sanitation systems had essentially eliminated waterborne diseases such as typhoid, dysentery, cholera,

Acknowledgments: The author would like to thank Mr. Ahmet Ozman and the team at Black & Veatch for their work on the *National Capital Region (NCR) Water Supply and Distribution System Redundancy Study: Project No. 188286*, which is the basis for this chapter's case study. The author would also like to thank the partner water utilities in the metropolitan Washington region for supporting and participating in the study.

1. World Bank Group, "Earth's Water," Open Learning Campus (website), accessed September 28, 2021, https://olc.worldbank.org/sites/default/files/sco/E7B1C4DE-C187-5EDB-3EF2-897802DEA3BF/Nasa/chapter1.html.

2. US Bureau of Reclamation (USBR), "Water Facts – Worldwide Water Supply," USBR (website), accessed November 22, 2021, https://www.usbr.gov/mp/arwec/water-facts-ww-water-sup.html.

3. US Geological Survey (USGS), "Water Q&A: How Much Water Do I Use at Home Each Day?," USGS (website), June 20, 2019, https://water.usgs.gov/edu/qa-home-percapita.html.

4. Centers for Disease Control (CDC), "MMWR Weekly: Ten Great Public Health Achievements—United States, 1900–1999," CDC (website), April 2, 1999, https://www.cdc.gov/mmwr/preview/mmwrhtml/00056796.htm.

and Hepatitis A.[5] Even with these improvements, waterborne diseases are still responsible for hundreds of thousands of deaths each year due to lack of clean water and sanitation. This is a stark reminder that when disaster strikes, one of the most urgent needs facing communities is water for drinking, sanitation, hygiene, and firefighting.

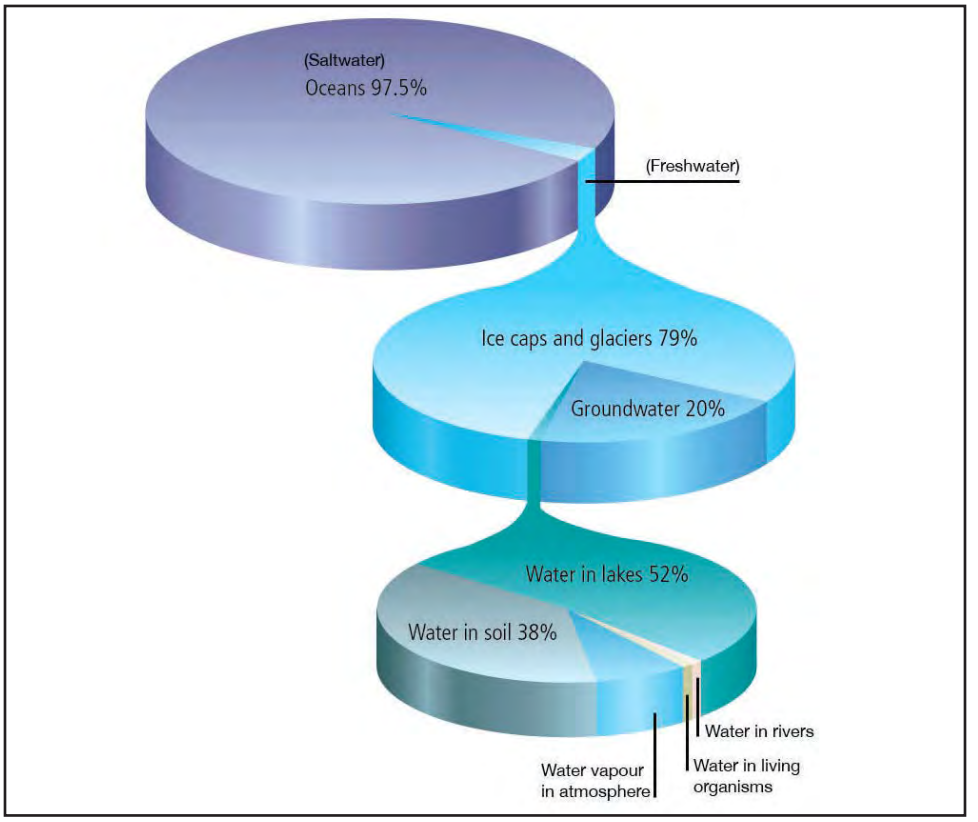


Figure 8-1. Earth's water
(Graphic by World Bank Group)

Since all other critical infrastructure sectors depend upon the water sector, the US Department of Homeland Security (DHS) designated water as a lifeline sector. See chapter 1 for the relationship between lifeline sectors and other critical infrastructure sectors.[6] A study published by the National Infrastructure Advisory Council in 2016 contains key findings

5. National Academy of Engineering, "Water Supply and Distribution Timeline," Greatest Engineering Achievements of the 20th Century (website), n.d., accessed September 28, 2021, http://www.greatachievements.org/?id=3610.

6. Department of Homeland Security (DHS), *NIPP 2013: Partnering for Critical Infrastructure Security and Resilience* (Washington, DC: DHS, 2013), 6, https://www.cisa.gov/sites/default/files/publications/national-infrastructure-protection-plan-2013-508.pdf.

about the importance of water infrastructure, especially how the loss of water services can impact other critical infrastructure and cause various types of disruption. In particular, the study indicates that "among surveyed critical infrastructure that depend upon water for core operations, services are degraded 50 percent or more within eight hours" after losing water or wastewater service.[7] Figure 8-2 portrays this relationship between the loss of water and the degradation of services by various sectors.[8] Despite significant dependency on the water sector, many critical infrastructure owners and operators do not have adequate plans for alternative sources of water or wastewater service. See chapter 12 for greater detail on the nature of dependencies and interdependencies, and their importance in developing risk management strategies.

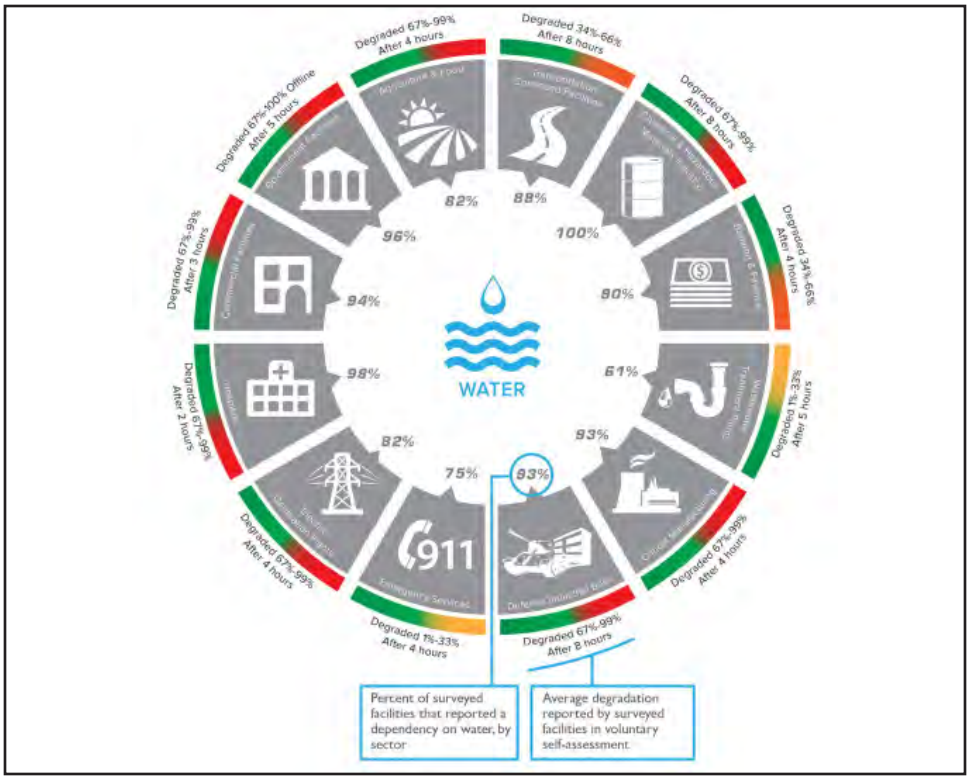


Figure 8-2. Sector dependency on water

(Graphic by US National Infrastructure Advisory Council. The information provided in the graphic is based on a limited sample of 2,661 voluntary facility assessments conducted between January 2011 and April 2014, DHS Sector Resilience Report, 2014)

7. National Infrastructure Advisory Council (NIAC), *Water Sector Resilience Final Report and Recommendations* (Washington, DC: NIAC, June 2016), 1, https://www.cisa.gov/publication/niac-water-sector-resilience-final-report.

8. NIAC, *Water Sector Resilience*, 2.

In the United States, more than 90 percent of the population relies on one of about 153,000 public water systems to provide clean water to homes and businesses. Of these systems, most people (over 80 percent) rely on a few large or very large water systems—such as rivers, lakes, reservoirs, or groundwater, which is treated before being distributed to customers—for their drinking water.[9] The remainder of the American population relies on private groundwater wells for drinking water. Once the water is used, it is discharged to a wastewater treatment system. For roughly 75 percent of Americans, this means using a public wastewater collection and treatment system. Of those served by a public wastewater system, most receive service from one of 382 large and very large wastewater treatment systems serving large urban areas of more than 100,000 people, or from one of the roughly 16,000 smaller systems serving less-populated cities. The remaining 25 percent of people depend on individual, onsite or small community cluster (septic) systems to treat their wastewater.[10] See figure 8-3 for a depiction of the water cycle and the relationship between water source, distribution and treatment.[11]

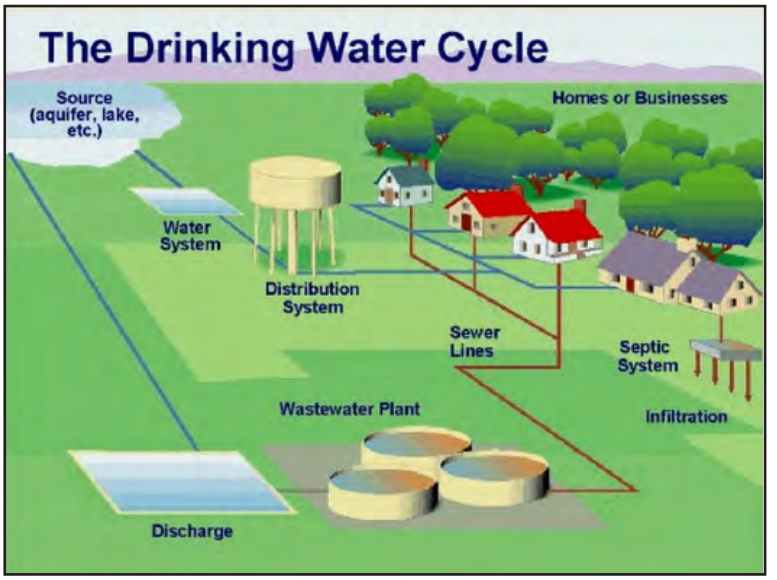


Figure 8-3. Drinking water cycle
(Diagram by University of Florida IFAS Extension)

9. Environmental Protection Agency (EPA), *Fiscal Year 2010: Drinking Water and Ground Water Statistics* (Washington, DC: EPA, 2011), 4.

10. DHS and EPA, *2015 Water and Wastewater Sector-Specific Plan* (Washington, DC: DHS, 2015), 5–6, https://www.cisa.gov/sites/default/files/publications/nipp-ssp-water-2015-508.pdf.

11. Amy L. Shober and Alexander J. Reisinger, *Drinking Water Source Protection in the Tampa Bay Region: A Guide for Homeowners* (Gainesville: University of Florida IFAS Extension, August 29, 2021), 2, https://edis.ifas.ufl.edu/pdf/SS/SS493/SS493-D3qinldqtf.pdf.

To put this relationship into perspective on a local scale, over five million people in the metropolitan Washington region are served by 18 major wastewater treatment plants, 13 drinking water suppliers, and 27 distributors. These water utilities own, operate, and maintain about 16,000 miles of sewer pipes, 14,500 miles of drinking water mains, and more than 114,000 fire hydrants.[12] One facility, the Blue Plains Advanced Wastewater Treatment plant, serves more than two million people and treats, on average, about 300 million gallons of wastewater per day.[13]

Risks and Threats to the Water Sector

Owners and operators of drinking water and wastewater systems must manage their assets in a manner that effectively addresses their risk profile. Factors such as utility location, size, assets, and specific risks will determine the risk management priorities for each utility. Common risks and threats to the water sector can be divided into two broad categories. The first category is intentional threats, such as cyberattacks; destruction (physical or via cyberattack) of parts of a system; contamination of drinking source water using various chemicals, toxins, microbes, or radioactive compounds; and contamination of treated water in the distribution system. The second category, natural hazards or unintentional events, includes extreme weather and climate change; aging infrastructure; accidental contamination of drinking source water, like oil or toxic material spills and combined sewer overflow; and accidental contamination of treated drinking water, such as a treatment plant process malfunction or human error, failed backflow prevention device, and loss of pressure causing inflow of contaminated water.

The concept of an intentional attack or terrorist threat to water supplies is not new. Shortly after the attack on Pearl Harbor in 1941, Federal Bureau of Investigation (FBI) Director J. Edgar Hoover wrote that "among public utilities, water supply facilities offer a particularly vulnerable point of attack to the foreign agent, due to the strategic position they occupy in keeping the wheels of industry turning and in preserving the health and morale of the American populace."[14] In January 2002, not long after the 9/11 terror attacks, open source reporting and an alert from the FBI's National

12. Metropolitan Washington Council of Governments (MWCOG), *State of the Region: Infrastructure Repor*t (Washington, DC: MWCOG, 2015), 24.

13. "The Largest Advanced Wastewater Treatment Plant in the World," DC Water (website), n.d., accessed September 28, 2021, https://www.dcwater.com/blue-plains.

14. John Edgar Hoover, "Water Supply Facilities and National Defense," *Journal of the American Water Works Association* 33, no. 11 (1941): 1862, http://www.jstor.org/stable/41232575.

Infrastructure Protection Center indicated that al-Qaeda members had "sought information on water supply and wastewater management practices" and had an "interest in insecticides and pest control products," possibly indicating an interest in water supply contamination.[15] More recently, a global analysis carried out in 2018 by researchers at Florida International University found that water-related terrorism had increased 263 percent from 1970–2016, with 68 percent of those incidents occurring in the post-9/11 era.[16] Table 8-1 highlights several of these more recent attacks, most of which are attempts to poison water supplies with pesticides that are available commercially and relatively easy to obtain.[17]

Table 8-1. Recently documented water attacks

Attacks on Water Resources	
Year	**Description**
2002	Four men of the Salafist Group for Preaching and Combat are arrested in Rome in possession of chemicals, false papers, and detailed plans for the water supply network in the zone of the Embassy of the United States.
2002	Two al-Qaeda agents arrested in Denver with plans to poison water resources.
2004	The FBI and DHS warn that terrorists are trying to recruit employees from water treatment plants as part of a project to poison drinking water.
2006	A water tank in Tring, England, is deliberately contaminated with herbicide.
2006	Strychnine (a pesticide) is intentionally released into a Danish artificial lake.
2007	201 people in China die after using water that had been intentionally contaminated with fluoroacetamide (a pesticide).
2008	A man was arrested in Varney, Virginia, in possession of two vials of cyanide and attempting to poison the water supply system.
2008	The water supply system of a Burmese refugee camp in Thailand (with a population of 30,000) is intentionally poisoned with herbicide.
2009	In the Philippines, the group *Frente Moro de Liberación Islámica* poisons water sources used by government soldiers and the general population.

15. Steven J. Duranceau, C. David Plavacan, Rick Hahn, and William J. Ackerman, "Al Qaeda and Your Water," *2002 Florida Section American Water Works Association Conference Proceedings* (Palm Harbor, FL, November 2002), 4, https://ssem.eku.edu/sites/ssem.eku.edu/files/files/HAHN%20-%20AWWA%20Paper%20Al-Qaeda%20and%20Your%20Water-Final.pdf.

16. Jennifer Veilleux and Shlomi Dinar, "New Global Analysis Finds Water-Related Terrorism Is on the Rise," NewSecurityBeat (website), May 8, 2018, https://www.newsecuritybeat.org/2018/05/global-analysis-finds-water-related-terrorism-rise/.

17. "Locken: Understanding Terrorist Threat to Our Water Sector," Security News Desk (website), January 9, 2017, https://securitynewsdesk.com/locken-understanding-terrorist-threat-water-sector/; and Agencies in Pristina, "Kosovo Cuts Pristina Water Supply over Alleged Isis Plot to Poison Reservoir," *Guardian (website)*, July 11, 2015, https://www.theguardian.com/world/2015/jul/11/kosovo-cuts-pristina-water-supply-over-alleged-isis-plot-to-poison-reservoir.

Attacks on Water Resources (continued)	
Year	**Description**
2010	Maoist rebels poison a pond in the Kashmir region, which the Central Reserve Police Force used as a source of drinking water.
2010	In England, two neo-Nazis are convicted of several counts of terrorism, including castor-making and conspiracy with Serbian Nazis to poison water resources used by Muslims.
2011	Documents seized during the raid to kill Osama bin Laden reveal plans for poisoning water resources.
2011	In Cadiz, Spain, officials thwart a conspiracy to poison water resources in response to the death of Osama bin Laden.
2012	Two 5,000-liter drinking water tanks in Australia are deliberately poisoned with Diuron (a herbicide).
2012	Hundreds of girls in an Afghan school become sick after deliberate poisoning of the water supply system.
2015	Five people linked to the Islamic State (Da'esh) arrested in Kosovo for allegedly planning to poison a reservoir. Authorities in Kosovo cut off the water supply to tens of thousands of people in Pristina to test the water for contaminants.

It is highly unlikely that an attack against drinking source water (such as a river, lake, or reservoir) would be successful due to the large volume of water, dilution of the contaminants, and their removal by water treatment. An attack in the drinking water distribution system, however, has a higher probability of causing harm. In one such attack in 1984, for example, members of a religious cult contaminated a water supply tank in The Dalles, Oregon, leading to 750 confirmed cases of salmonella.[18] While the probability of a high-consequence attack might be low, the terrorist threat to water systems is real.

As do all other sectors, the water sector faces a growing cyber threat to its industrial control systems, business systems, e-mail, and information technology infrastructure. A broad array of threat actors—including nation states, extremists, and criminals—take advantage of expanding reliance on and connectivity of information technology systems to launch attacks and achieve their objectives. See chapters 3–5 for an overview of various cyber threats and actors. Specific to the water sector, a February 2021 cyberattack on a water treatment plant in Oldsmar, Florida, attempted to increase levels of sodium hydroxide, a chemical used to treat the water. While the intrusion

18. Peter H. Gleick, "Water and Terrorism," *Water Policy 8* (2006): 487, https://pacinst.org/wp-content/uploads/2013/04/water_and_terrorism_2006.pdf.

was quickly detected and chemical levels returned to normal, this example illustrates how even unsophisticated hackers can access water sector software and control systems that are connected to the Internet, putting water sources and people at risk.[19]

In the category of unintentional threats, the metropolitan Washington region, like other areas of the world, is experiencing the impacts of a changing climate. According to the Metropolitan Washington 2030 Climate and Energy Action Plan, "temperatures and the water surface level in the Potomac River have been rising and will continue to rise. Extreme weather events and increases in the number of extreme heat and cold days will increase risks to health, energy usage patterns, plant and animal habitats, and infrastructure."[20] Drinking and wastewater systems will be affected by these changes and will need to adjust management strategies to protect their infrastructure from damage while continuing to meet water quantity and quality requirements.

In many NATO countries, large portions of water infrastructure were built in the early twentieth century. Much of this aging infrastructure has reached or will soon reach the end of its useful lifespan and need to be replaced or upgraded. In the metropolitan Washington region, one such example is WSSC Water, which has served Maryland residents in Montgomery and Prince George's Counties since 1918. More than 40 percent of WSSC Water's 11,000 miles of underground pipes are over 50 years old. These older pipes have reached the end of their service lifespan, but replacing these pipes—at an average replacement cost of $1.4 million per mile—is a very expensive and lengthy project.[21] Replacing aging infrastructure is a high-priority issue that must be weighed against other potential threats.

Water Sector Approaches to Resilience Planning

There are numerous resources and approaches to resilience planning in the water sector. In the United States, the Water Sector-Specific Plan—an annex to the National Infrastructure Protection Plan, published in 2007, 2010, and 2015 by the DHS and the Environmental Protection

19. Andy Greenberg, "A Hacker Tried to Poison a Florida City's Water Supply, Officials Say," Wired (website), February 8, 2021, https://www.wired.com/story/oldsmar-florida-water-utility-hack/.

20. MWCOG, *Metropolitan Washington 2030 Climate and Energy Action Plan* (Washington, DC: MWCOG, 2020), 7.

21. "Aging Infrastructure," WSSC Water (website), March 23, 2021, https://www.wsscwater.com/what-we-do/major-projects/pipes-and-infrastructure-improvements-and-maintenance/aging.

Agency (EPA)—addresses risk-based critical infrastructure protection implementation strategies for drinking water and wastewater utilities. It describes processes and activities to increase resilience in the sector and prepare to prevent, detect, respond to, and recover from hazards. The plan follows a risk management framework that sets goals and objectives, identifies infrastructure to protect, assesses and analyzes risks to that infrastructure, implements risk management activities, and measures the effectiveness of these actions. The plan also sets out four goals for the water sector: (1) sustain protection of public health and the environment; (2) recognize and reduce risk; (3) maintain a resilient infrastructure; and (4) increase communication, outreach, and public confidence.[22]

Several guidance documents and standards for the water sector have been developed to better understand and reduce risks to water and wastewater services. On its website, the American Water Works Association has some state-of-the-art standards that are available for a small fee. Three of these standards, which NATO Allies and partners may also find useful, are: (1) Risk Analysis and Management for Critical Asset Protection Standard for Risk and Resilience Management of Water and Wastewater Systems (AWWA J100); (2) Security Practices for Operation and Management (AWWA G430); and (3) Emergency Preparedness Practices (AWWA G440).[23] Likewise, the EPA has also developed some risk assessment tools for drinking water and wastewater utilities of all sizes, which can help identify the highest risks to critical operations and identify cost-effective solutions to reduce those risks. Two useful tools, available on the EPA's website, are the Vulnerability Self-Assessment Tool and the Climate Resilience Evaluation and Awareness Tool Risk Assessment Application for Water Utilities.[24]

In the United Kingdom, events and risks associated with flooding, drought, extreme cold spells, and climate change over the past decade stressed water companies' ability to reliably provide water and wastewater services and have, as a result, increased the amount of attention paid to resilience. Consequently, the English and Welsh governments gave Water Services Regulation Authority (Ofwat) the responsibility to address resilience

22. DHS, *NIPP 2013*, 15.

23. See "American Water Works Association (AWWA) Publications Catalog," AWAA (website), n.d., accessed November 9, 2021, https://engage.awwa.org/PersonifyEbusiness/.

24. See EPA, "Conduct a Drinking Water or Wastewater Utility Risk Assessment," EPA (website), n.d., November 9, 2021, https://www.epa.gov/waterriskassessment/conduct-drinking-water-or-wastewater-utility-risk-assessment; and "Climate Resilience Evaluation and Awareness Tool (CREAT) Risk Assessment Application for Water Utilities," EPA (website), n.d., accessed November 9, 2021, https://www.epa.gov/crwu/climate-resilience-evaluation-and-awareness-tool-creat-risk-assessment-application-water.

in the regulation of the water sector. The 2014 Water Act required action to "further the resilience objective," and Ofwat developed a policy response to define requirements for water companies to improve the resilience of water and wastewater services using the concept of "resilience in the round." Ofwat used a holistic concept of resilience—the ability to prevent, cope with, and recover from disruptions of all kinds—that focuses on three key dimensions of resilience in an organization: corporate, financial and operational.[25] The corporate dimension is an organization's governance, accountability, and assurance processes, and its ability to anticipate trends and variability in its business operations. The financial aspect considers the impact of an organization's funds and assets, while operational resilience focuses on an organization's infrastructure and the ability and skills to run that infrastructure. Finally, the UK's Water Industry Research organization published its Good Practice Guide, which provides the water companies with planning guidelines and outlines a framework for risk screening, resilience assessment, resilience planning, and implementation of resilience solutions.[26]

This section provided a summary of the best practice components of the US and UK frameworks for resilience planning in the water sector. In the next section—focused on a case study of the US National Capital Region (NCR) water sector—the Metropolitan Washington Council of Governments (COG) adapted and employed several of these best practices in its evaluation of the water supply system's resilience. The COG first characterized the system, which involved describing and understanding the physical water system infrastructure in the NCR, evaluating its performance against recent risk events, and determining a risk assessment approach.

The next step, risk screening and assessment, included screening and assessment of hazards that pose a risk to service, considering how critical assets and systems respond to these hazards, assessing the potential level of service disruption, and estimating potential losses due to service disruption. Then, the COG conducted a detailed resilience assessment to evaluate the existing (baseline) level of resilience, identify and evaluate potential initiatives for increasing resilience, and determine alternatives using a whole life cycle cost-benefit analysis. The final step, implementation of resilience solutions, aimed at developing and implementing improvement initiatives to address resistance, reliability, redundancy, response, and recovery.

25. Ofwat, *Resilience in the Round: Building Resilience for the Future* (Birmingham, UK: Ofwat,2017): 3–4, https://www.ofwat.gov.uk/wp-content/uploads/2017/09/Resilience-in-the-Round-report.pdf.

26. UK Water Industry Research (UKWIR), *Resilience Planning: Good Practice Guide—Summary Report* (London: UKWIR, April 22, 2013), 1, https://ukwir.org/eng/reports/13-RG-06-2/66806/Resilience-Planning-Good-Practice-Guide--Summary-Report.

Metropolitan Washington Region Case Study

Background and Goals

The metropolitan Washington region is a diverse and dynamic area centered on the nation's capital in the District of Columbia, and surrounded by counties in suburban Maryland and northern Virginia. It is home to more than 5.5 million people and one of the nation's largest economies, with more than one million new residents and jobs forecasted between now and 2045. In a highly urbanized area like metropolitan Washington, a water emergency has the potential for significant regional disruption because there is limited capability to transfer potable water across the Potomac River or to areas where shortfalls might occur due to the segmented nature of the region's water systems. Depending on the scenario, the consequences of such a regional water outage could be the loss of water service from days to more than a month, while the estimated direct and economic impacts could be several billion dollars.

Since 2007, the COG, working with water utilities across the region, carried out three studies on water supply and distribution system redundancy. The main purpose of those studies was to evaluate the regional water supply system's ability to withstand emergencies and identify potential engineering improvements to increase the overall reliability of the system. The most recent of these studies, completed in 2016, is the focus of this section.

The 2016 study used a risk-based methodology and employed industry best practices to evaluate potential failure scenarios, their impacts on water service, societal and economic consequences, and potential mitigating initiatives. Carried out by a team from Black & Veatch, a leading water engineering firm, this project evaluated potential improvements to enhance regional water system resilience and assessed the reliability and safety of drinking water from source to tap from a regional perspective. Working with COG staff and other project participants, Black & Veatch scheduled a series of workshops from June 2015–March 2016 with water utilities in the region. The workshops laid the foundation for a successful study by: (1) establishing project goals and working relationships with key water utilities; (2) identifying the level of service goal to be achieved under emergency conditions; (3) describing failure events that could result in widespread water supply outages; (4) quantifying the likelihood of occurrence of failure events, duration of outages, and number of customers affected; (5) identifying infrastructure improvements to mitigate water supply outages in potential failure scenarios; (6) assessing and prioritizing improvement options with respect to return on investment and benefits

to regional resilience; and (7) reaching consensus on an improvement plan that optimizes risk reduction with cost and level of service performance.

To help maintain consistency and common understanding, the study uses these important terms and definitions:

- *Level of service* (LOS): a set of objectives that define the service performance to be delivered
- *LOS impact*: failure to meet LOS goals for some number of people for some duration because of a failure scenario and associated asset failures
- *Event*: an occurrence, such as a contamination, fire, or catastrophic asset failure, that can act as a root cause and create a LOS impact
- *Failure scenario*: the chain of incidents that occur due to an event occurrence that creates or exacerbates a LOS impact
- *Asset failure*: specific water system assets that fail as part of the failure scenario
- *Direct cost*: the immediate financial impact of a failure scenario, including the cost to repair or replace equipment
- *Economic cost*: collateral damage and costs not directly caused by the failure scenario
- *Financial cost*: the combination of direct and economic costs
- *Improvement initiatives*: capital projects that can decrease the likelihood or consequence of a failure scenario

Risk Assessment and Modeling

A key element in resilience analysis is risk analysis. See chapter 13 for an overview of risk assessment and management considerations and strategies. Following common practice for evaluating risk, this study calculated risk by multiplying the likelihood of an event occurring by the consequence of its occurrence. In the initial workshops, Black & Veatch worked with participating water utilities to identify various events that might result in an extended loss of water service and estimate the likelihood of such events occurring as well as their consequences. Quantifying risk and resilience for many potential initiatives and portfolios requires a defined process to ensure the study alternatives are evaluated in a consistent manner.

To that end, figure 8-4 illustrates the study approach at the macro level while the rest of this section offers a more detailed description of each step of this process.[27]

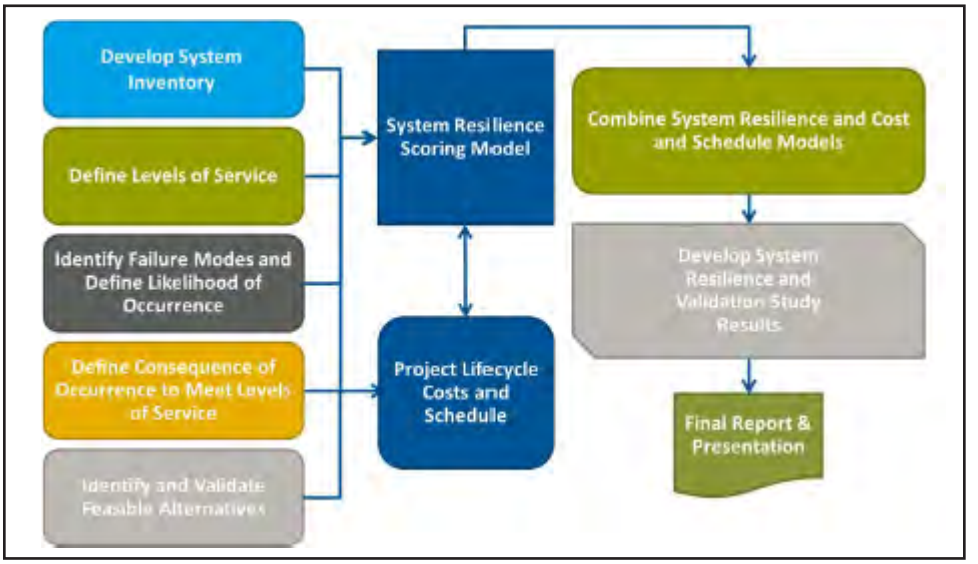


Figure 8-4. Process for the NCR water system study
(Diagram by Black & Veatch)

Step 1: Develop System Inventory

Evaluating the regional system's resilience required an understanding of the NCR's existing individual water systems. Prior to carrying out risk assessment and modeling work, it was necessary to inventory the existing critical regional water systems in the NCR, such as water treatment plants, raw water sources, water storage, pumping facilities, and key interconnections. Population and water demand projections were also obtained, as were any available previous studies and plans by participating water utilities. The region's water supply system effectively serves the demands in the NCR under normal operating conditions. Regional agreements and operational coordination between water utilities concerning Potomac River raw water usage, reserve raw water, treated water capacity, emergency preparedness, and mutual aid further enhances the reliability of the system. Population forecasts

27. Black & Veatch, *National Capital Region (NCR) Water Supply and Distribution System Redundancy Study: Project No. 188286* (Washington, DC: MWCOG, July 2016).

prepared by COG estimated modest but continued growth through 2040 in the entire region.[28]

Based on these population forecasts, the 2015 Washington Metropolitan Area Water Supply Study assessed the ability of the NCR's current water supply system to meet forecasted demands through 2040. The study concluded that—given conditions similar to severe historic droughts and assuming no impact from climate change—the "current water supply system will experience considerable stress, with mandatory water use restrictions required" in the Washington metropolitan area by 2035.[29] The region has studied several raw water storage and transmission improvements (such as conversion of quarries to off-river reservoirs) to provide additional protection against drought risks. These options will be developed and refined further as drought risk and impacts of climate change are better understood.

As for water treatment capacity, the region has adequate treatment capacity to meet maximum day demands through the year 2040. Due to the segmented nature of different water systems, there is limited capability to transfer treated water across the Potomac River and to areas where shortfalls might occur. For example, in the event of a water treatment plant outage, there could be limited ability to transfer water from other treatments plants to meet demand in the affected area.

Step 2: Define Levels of Service (LOS)

A system-level resilience analysis requires at least one baseline metric to assess the performance of the existing system and compare it to future, alternative systems. For this study of regional water system redundancy, the baseline metric was the ability of customers to receive potable water when disruptive events occur. Specifically, through the workshop process, the ability to supply winter average demands during emergency events was identified as the LOS.

To determine the LOS, estimates of 2040 winter average demands were calculated for each water utility. This LOS is consistent in principle with regional agreements such as the Potomac River Low Flow Allocation

28. MWCOG, *Summary of Intermediate Population Forecasts, Final Round 9.0 Cooperative Forecasts* (Washington, DC: MWCOG, November 9, 2016), https://www.mwcog.org/documents/2018/10/17/cooperative-forecasts-employment-population-and-household-forecasts-by-transportation-analysis-zone-cooperative-forecast-demographics-housing-population/.

29. S. N. Ahmed, K. R. Bencala, and C. L. Schultz, *2015 Washington Metropolitan Area Water Supply Study: Demand and Resource Availability Forecast for the Year 2040* (Rockville, MD: Interstate Commission on the Potomac River Basin, August 2015), xiv–xv, https://www.potomacriver.org/wp-content/uploads/2015/08/ICP15-04a_Ahmed.pdf.

Agreement and the Metropolitan Washington Water Supply and Drought Awareness Response Plan for the Potomac River System.[30] These agreements call for allocation of water withdrawals in proportion to each utility's average winter water production when restrictions are in effect.

Step 3: Identify Failure Modes

Once the LOS was defined, a workshop was held with participating water utilities to identify the types of major events that could impact these levels for large numbers of customers. Numerous events were considered and screened, including drought, earthquakes, river contamination, derecho winds, cyberattacks, power outages, and ice storms. Terrorist attacks were not singled out as an event, though several of the events that were considered (such as cyberattacks, river contamination, and power outages) could be caused by terrorism. Since this study focused on region-wide resilience, the failure scenarios evaluated in this study must either cause water outages across the region for more than 24 hours or affect important interconnections between water utility systems, causing water outages in large sections of a distribution system for more than 24 hours. Based on these criteria, the study group selected 10 failure events to evaluate, as outlined in table 8-2.[31]

Some failure scenarios—including events such as a regional power outage, windstorm, ice storm, cold weather events, and cyberattacks—were screened from the original list based on the workshops and subsequent discussions with relevant utilities. Ultimately, the selection of failure scenarios to be evaluated depends upon many factors such as the water utilities involved, existing investments in system redundancy, and the current threat environment.

30. *Potomac River Low Flow Allocation Agreement* (December 15, 2017), 11, https://www.potomacriver.org/wp-content/uploads/2018/02/LFAA-Annotated_2_22_2018.pdf; and MWCOG, *Metropolitan Washington Water Supply and Drought Awareness Response Plan: Potomac River System* (Washington, DC: MWCOG, June 7, 2000): 3–5, https://www.mwcog.org/file.aspx?D=%2Fy%2BiVOLXyZ%2B5oVxyNNqz5KFEhqK%2FQWPDwi26CumRG-%3D&A=%2B14Zj%2B5VOXzO6Z7cOTabI1TKNSn8nDbIWjV0E8pIRdk%3D.

31. Black & Veatch, *Project No. 188286*.

Table 8-2. Events and failure scenarios

Event	Failure Scenarios
Water main break	Major water main break across the Potomac River
Contamination in the Potomac River	Oil spill in the Potomac River—affects both banks
	Oil spill in the Potomac River—affects west bank
	Oil spill in the Potomac River—affects east bank
	Chemical spill in the Potomac River—affects both banks
Fire	Fire at a water treatment plant—west side of river
	Fire at a water treatment plant—east side of river
Airplane crash	Airplane crash into a water treatment plant—west side of river
	Airplane crash into a water treatment plant—east side of river
Reservoir contamination	Contamination of a drinking water reservoir

Step 4: Define Likelihood of Occurrence (LOO)

To carry out a resilience and risk analysis, the likelihood of selected failure events occurring must be determined. It is not possible to predict future events with certainty, so one workshop focused on determining reasonable estimates of the LOO based on the experiences of the water utilities participating in the study and others around the industry. For resilience analysis purposes, the most important aspect is to estimate the LOO for the various events relative to each other and calculate the magnitude of the differences. For example, the utilities participating in this study experienced oil spills in the Potomac River over the past few decades, but a plane crash into a water treatment plant has not occurred and is, intuitively, far less likely. Consequently, Potomac River contamination events from oil were not estimated to be similar in likelihood to a plane crashing into a treatment plant.

This study established five levels of LOO that provided large enough intervals for resilience modeling. The LOO in this study was expressed as having a single occurrence over a specific period, such as once every 30 years. Table 8-3 shows the five levels of LOO considered in the study, their definitions, and an additional description for the frequencies considered.[32]

For use in the resilience modeling approach, the study team worked with COG to associate a specific LOO with each failure scenario described above. All LOO levels were considered individually for each failure scenario,

32. Black & Veatch, *Project No. 188286*.

but only the moderate (once every 30 years) and very low (once every 250 years) levels emerged for the final frequencies.

Table 8-3. Five levels of likelihood of occurrence (LOO)

Rating	Definition	Frequency
Very high	There is direct evidence or substantial indirect evidence to suggest the failure scenario has initiated and/or is likely to occur.	Once per year
High	The failure scenario is known to exist, indirect evidence suggests it is plausible and key evidence is weighted more toward likely than unlikely.	Once every 10 years
Moderate	The failure scenario is known to exist, indirect evidence suggests it is plausible and key evidence is weighted more toward unlikely than likely.	Once every 30 years
Low	The failure scenario cannot be ruled out, but there is no compelling evidence to suggest it has occurred or that a condition or flaw exist that could lead to its development.	Once every 100 years
Very low	Several events must occur concurrently or in a series to trigger failure. Most, if not all, the events are very unlikely to ever occur.	Once every 250 years

Step 5: Define Consequence of Occurrence (COO) to Meet Level of Service

Evaluating risk requires combining likelihood of occurrence with its corresponding consequence. To develop the COO for the various failure scenarios in this study, the team evaluated the impact to the NCR system as it pertains to the average winter day demand, which was the LOS defined in step 2. For purposes of this study, if a failure scenario would cause a utility to be unable to supply potable water at average winter day demand levels, then the customer would be counted as experiencing an outage. Customer outages due to the types of events in this study can extend for several hours or days, and the study combined the number of customers experiencing an outage with the anticipated duration of the outage (in days) for each failure scenario. To quantify the COO, the study team multiplied the number of customers experiencing an outage (population) with the duration of the outage (days). The product, referred to as population outage days (POD), was calculated for each of the failure scenarios in the study.

To ensure consistency in the modeling over a 100-year period, the study relied on a key assumption: it used projected water demands and population estimates in 2040 for each of the utilities regardless of the year a potential failure scenario might occur. This assumption allowed for the use of a well-accepted set of population data to calculate outage impact

over the period modeled. It is important to note one simplification for this study, which was to evaluate the risks only to residential customers. Greater weights were not assigned to critical customers and facilities or outages that might impact mission assurance for military installations. An analysis that focuses on critical customers or mission assurance may result in a different prioritization of improvements.

The study also quantified risk by using a methodology that combines direct and indirect costs to calculate the risk associated with each failure scenario. Direct costs are the immediate financial impacts of a failure scenario (such as the cost to repair or replace equipment). From a modeling perspective, direct costs are not addressed by the initiatives in the study. For example, there are repair costs associated with a water main break, but they do not prevent the break. Instead, they serve to mitigate the number of POD after the failure scenario occurs. Indirect (economic) costs are the collateral damage for failure scenarios and do not include direct costs. In this study, the economic costs were calculated by applying a per capita, per day cost to the associated POD caused by a failure scenario. The study used an estimate of $114.38 per day as the total impact due to the loss of potable water service per capita.[33]

Step 6: Identify and Validate Feasible Alternatives

After defining the failure scenarios and establishing the likelihoods and consequences of occurrence, the study identified improvement initiatives that address water system risk posed by the failure scenarios. The alternatives identified were potential countermeasures that might mitigate or prevent the effects of a failure scenario. Alternatives were selected by analyzing the failure scenarios and identifying approaches to reduce POD. Discussions with participating water utilities, engineering expertise and knowledge of the NCR system all contributed to the development of suitable alternatives. The study group then reviewed various conceptual alternatives with participating utilities during a series of meetings. All potential improvements were presented and considered during an improvement alternatives workshop and during subsequent discussions with utilities to focus on and refine potential improvement options for further consideration.

For each failure scenario, the study team developed a spreadsheet-based water supply mass balance model and post-improvement conditions. This model was used to estimate the benefit of each potential improvement in terms of capacity provided and reduction in water supply shortfall in each demand area. The improvement benefit was then calculated as a reduction

33. Black & Veatch, *Project No. 188286*.

of POD from base case conditions. Types of improvements assessed included off-river water storage, improved water transmission, treated water transfer, and interconnections between water utilities. Estimates of costs for capital and operations and maintenance were calculated for each improvement option. Implementation schedules were also estimated, considering typical durations needed for planning, design, permitting, and construction for each improvement.

The system resilience model used in this study is a Microsoft Excel-based model, which utilizes a Monte Carlo analysis to calculate the risk of a scenario in which the NCR does not implement resilience measures during the study period. This scenario is the base case. As the risk and potential impacts to the LOS were better understood and quantified, potential infrastructure improvement options were developed to mitigate outages and the risk involved in the base case. These potential improvements were reviewed, screened, and developed with input from relevant water utilities. A total of 18 potential improvements were developed and evaluated for their benefits, in terms of reducing the number of POD for each failure scenario. The study team also developed cost estimates for construction, and operations and maintenance as well as implementation schedules considering time required for planning, design, and construction.

Results

The evaluation of the base case yielded a net present value cost of $37 billion in monetized risk in the existing system over the 100-year modeling period. The $37 billion value is comprised of the risk associated in the 10 failure events outlined earlier in Step 3 (see table 8-2): large water main failure ($0.11 billion), Potomac River contamination with oil ($20.1 billion), Potomac River contamination with oil on the east bank of the river ($1.31 billion), Potomac River contamination with chemicals ($9 billion), a fire at a major water treatment plant ($0.5 billion), and reservoir contamination ($5.56 billion). Events involving contamination of the Potomac River are responsible for a substantial amount of the total risk carried by the existing NCR water system.

Off-river water storage combined with raw water transfer or treated water transfer improvements were shown in this study to be effective risk mitigating initiatives. Each potential improvement was individually modeled to determine its benefit over a 100-year modeling period. Resulting cost-benefit ratios were developed and used to rank improvements. Improvements were then grouped into portfolios of improvements for selected

spend levels to achieve the maximum possible cost-benefit ratio. A total of 27 individual portfolios were defined and modeled covering a capital spend range from $1 million to roughly $330 million, as figure 8-5 illustrates.[34] As expected, model results indicate a reduction in overall risk with increasing spend levels.

The risk modeling results provide a basis for an objective evaluation of potential improvement options with respect to their life-cycle costs and resilience benefits to the region. The risk analysis showed a total monetized net present value risk of $37 billion over the 100-year modeling period and highlighted that Potomac River contamination events are responsible for a substantial amount of the total risk carried by the region's water systems. The study revealed that targeted raw water storage combined with raw water transfer and treated water transfer improvements—generally considered "no regrets" types of improvements—are the most effective risk-mitigating initiatives, while providing a greater volume of raw water storage is a top longer-range improvement.

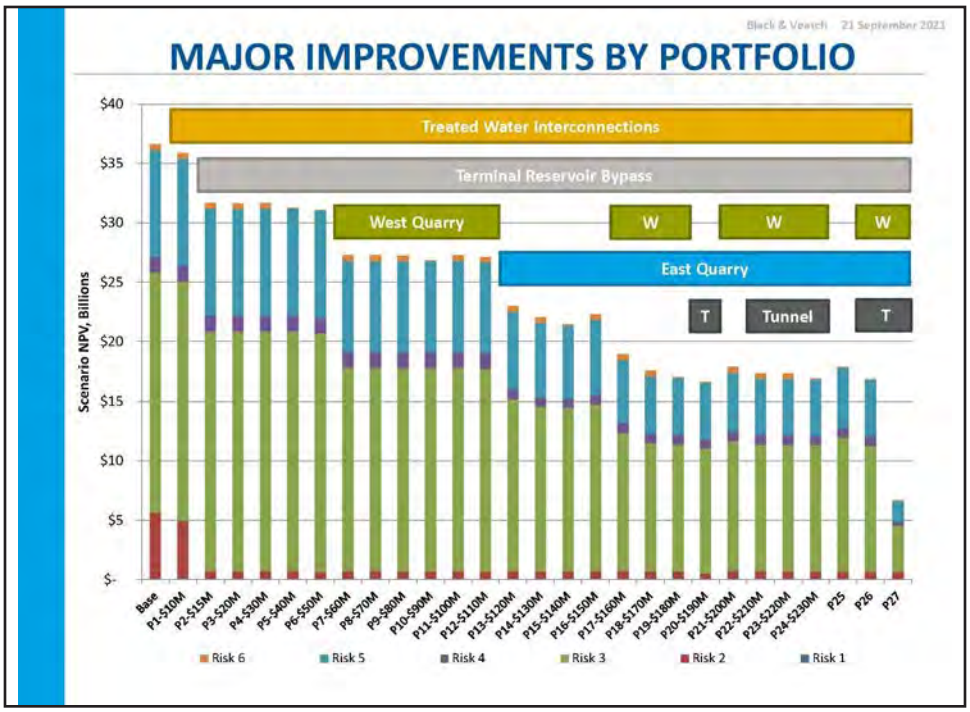


Figure 8-5. Major improvements by portfolio
(Source: Black & Veatch)

34. Black & Veatch, *Project No. 188286*.

While the case study of the water supply system in the NCR is unique to the Washington metropolitan area, it provides a framework and process for risk and resilience analysis that NATO Allies and partners can adapt and use to conduct similar studies of their water supply systems. The final section now provides recommendations for ways allies and partners can enhance security and resilience of local and national water infrastructure.

Recommendations and Actions for Consideration

Water plays a significant role in the lives of a nation's civilian population and in the ability of its armed forces to conduct the full range of military operations from peacetime to conflict. Daily conveniences (such as taking a shower, making coffee, brushing teeth, feeding a family, flushing the toilet, and washing laundry) are often taken for granted, but depend on water. The infrastructure needed to treat and convey drinking water and wastewater, much of it buried and out of sight, is essential to keeping businesses open and supporting other critical infrastructure sectors, including military installations and missions. When water services are significantly disrupted for several hours or longer, hospitals may have to close, hotels and restaurants cease operations, factories shut down, and military installations and the surrounding communities are disrupted. Water is a lifeline of the world's communities, and, as discussed throughout this chapter, it is important to take actions to protect water infrastructure from threats such as terrorism, cyber attacks, climate change, and more. To that end, there are four vital steps that NATO Allies and partners can take to make their water infrastructure more secure and resilient.

First, they can pursue regional collaboration and coordination with trusted partners to conduct water supply and wastewater treatment planning. Coordinated regional planning has many advantages, including the ability to share benefits, risks, and resource costs. Trusted relationships can also form the basis of a mutual aid network, helping water utilities to respond to and recover from emergencies.[35]

Second, they can conduct a condition assessment of water infrastructure using a risk-based approach. This approach allows utilities to inventory assets, identify high-risk assets, and determine which assets would have the highest consequence of failure in terms of the critical customers affected, number of customers affected, and other factors. For organizations facing

35. "Water/Wastewater Agency Response Network," AWWA (website), n.d., accessed September 28, 2021, https://www.awwa.org/Resources-Tools/Resource-Topics/Water-Wastewater-Agency-Response-Network.

budgetary constraints, this step can be a relatively low-cost approach that can be gradually expanded into a full asset management program in the future.[36]

Third, they can use a risk-based approach to balance risk reduction and cost. Every water utility or infrastructure owner or operator that depends on water has a unique risk profile. When assessing alternatives to mitigate risk, it is important to quantify the level of service needed in an emergency, the likelihood of failure, and the consequence of failure, and to look for synergies among mitigation alternatives.

Fourth, they can plan with a long-term vision in mind. Major capital improvements to water infrastructure are costly, sometimes in the range of hundreds of millions of dollars. Once constructed, a typical service life can be 50 years or longer. Consequently, investments in resilience measures made now should anticipate future conditions—to include growth in population, water demand, and threats from climate change or extreme weather—so the expected benefits and return on investment will be realized.

Strengthening the security and resilience of water infrastructure among NATO Allies and partners is a tremendous challenge. As the case study illustrates, using a risk-based approach to resilience planning can provide a solid foundation for setting priorities and determining the timing and funding levels of infrastructure investments to enhance an organization's or a region's ability to adapt and respond to water needs during emergencies.

36. Nelson Carriço and Bruno Ferreira, "Data and Information Systems Management for the Urban Water Infrastructure Condition Assessment," *Frontiers in Water (*website), July 5, 2021, https://doi.org/10.3389/frwa.2021.670550; and Ahmad Habibian, "Prioritizing Your Aging Water Infrastructure Needs through a Sytematic, Comprehensive Approach," Water Finance & Management (website), September 4, 2018, https://waterfm.com/condition-assessment-cornerstone-asset-management/.

— 9 —

Communications Resilience

Chris Anderson

Communications form the critical backbone of the modern world, connecting more people and more devices more completely than ever before. The benefits of this hyper-connected society drive ever-increasing reliance on secure, reliable, and resilient communications. Potential adversaries to the North Atlantic Treaty Organization certainly understand the importance of communications—those they seek to target and those they use themselves—so it is critical to fully understand the sector, the risks it faces, and the best ways to mitigate those risks. This chapter addresses these vital issues in four sections:

- An overview of the communications sector
- Threats to communications: a discussion of ways in which the integrity, availability, or confidentiality of communications systems may be degraded or compromised
- Case studies: key observations and lessons learned from several recent examples of incidents—natural and man-made, cyber and kinetic—that have targeted communications systems and related infrastructure
- Recommendations: suggestions for improving communications resilience against terrorist attacks and other threats

Communications Sector Overview

This overview provides a foundation from which to better understand the criticality of communications for national security and emergency preparedness, the current state of the sector and future trends, and common important characteristics of the sector and their implications for security and resilience.

Critical for National Security and Emergency Preparedness

Resilient and trustworthy communications are fundamental to national security and emergency preparedness. Communications, including public commercial networks, play a critical role in:

- National command and control, military operations, and the distribution of intelligence and warning
- Civil defense, law enforcement, and first-responder coordination
- Citizen preparedness and resiliency during crisis

Vital central government and military communication, including classified and unclassified information, flows over commercial communications channels. While tactical-level military communications still likely use fully noncommercial communications such as battlefield datalinks and tactical radios, most other federal and military traffic—including national command and control networks, intelligence collection and production, big-data analysis, automation, and multimedia voice and video connectivity—have high bandwidth and low latency requirements that are often best served by using commercial providers. In many cases, this vital government traffic flows alongside or even intermixes with civilian traffic through the same commercial fiber lines and data centers that support civilian communications.

At a more local level, civil defense, law enforcement, and other first responders also leverage a combination of self-managed communications systems and commercial networks to support operational coordination. Traditional first-responder communications have evolved from basic two-way land-mobile radios to include a wide range of radio frequency and fiber-optic voice and data connectivity. Even these land-mobile radio systems are now highly integrated systems that incorporate mesh networking and utilize fiber optics and satellite for backhaul of data and interconnection to other networks and data systems.

The United States is deploying and enhancing FirstNet, a National Public Safety Broadband Network designed to be a high-speed, nationwide wireless broadband network dedicated to providing voice, video, and data services in support of public safety. FirstNet is intended to be a reliable, highly secure, and interoperable communications network for public safety agencies and first responders, allowing them to get more information quickly and helping them to make faster and better decisions. Commercial carriers are offering similar "public safety grade" services to police, fire, rescue, and response agencies across the United States and among many NATO member states. As first-responder mission requirements increasingly incorporate data from video cameras, autonomous vehicles, and Internet of Things (IoT) sensors, the bandwidth needs and latency requirements for responders will likely continue to drive demand for ever more capable and complex networks to support them.

Communications are also essential for individual citizen preparedness. During times of emergency, citizens must receive timely and accurate news along with instructions from response officials on measures they should be taking such as evacuation, shelter-in-place, and "be on the lookout for . . ." notices. A primary means for citizens to receive this information is through broadcast media such as radio and television (TV). Broadcast is useful in its simplicity in providing a one-to-many communications path. The relatively low power demands for radio make it an excellent medium for survivable, even self-powered, communication to a mass audience.

Cellular communications have supplanted landline telephony in most market areas, but both are essential tools for citizens to check in on family members and reach out for help when necessary. Even as landline telephony has waned, high-bandwidth connectivity to the home has exploded, bringing a host of on-demand content—including access to a vast amount of public safety information—to consumers across the globe.

Both broadcast and cellular communications are leveraged in emergency alerting systems that provide reliable, survivable communications pathways for citizen critical emergency information. In the United States, the Wireless Emergency Alert system leverages cellular delivery of alert information while the Emergency Alert System provides alerts via radio and TV broadcast. Starting in June 2022, Article 110 of the European Electronic Communications Code will make it mandatory for all European Union member states to deploy

a public-warning system using telephone networks to alert everyone located in a specific area of an ongoing crisis or impending disaster.[1]

Common Sector Characteristics

Modern communications networks are highly meshed, extremely resilient multi-mode. Equipment failures or damaged fiber-optic cables are generally rerouted easily with little to no impact to end users. The weakest link in a communications network is often the "last mile" to a given consumer or enterprise. This final stretch, whether provided by copper lines, fiber-optic cable, or through the airwaves, is often a single-path connection, particularly in rural and less densely populated regions, and therefore does not provide resilience through redundancy like more dense environments or specifically designed survivable communications architectures.

The last several decades have seen a shift away from most analog communications in favor of packet-based communication technologies. Most communication today, even simple voice telephony, is digitized, packetized, and sent over digital data networks. This transition happened in wired voice communications as Voice over Internet Protocol (VoIP) phones became commonplace by the early 2000s. In the cellular space, 3G cellular services brought an initial expansion of data connectivity to mobile devices, with a transition to VoIP for core cellular voice calls arriving in conjunction with the transition to 4G. The digital transition has also revolutionized TV and radio broadcasting. HD radio standards allow up to four stations of content to share a single radio channel while still delivering crystal clear sound, while digital TV has enabled the broadcast of high-definition signals over the air. The next generation TV standard, Advanced Television Systems Committee 3.0, will allow for over-the-top datacasting, enabling the ability to provide advanced emergency alerting that includes maps, graphics, hyperlinks, and other data to better convey emergency situational awareness and direct specific community responses.

In most countries, the wide array of industry segments and multiple strong companies means a highly competitive communications marketplace where rapid innovation is leveraged to competitive advantage. At the same time, network providers are often mutually dependent on each other's infrastructure in colocation spaces, for backhaul, and even last-mile connectivity to customers, so coordination among providers is regular and routine.

1. *EU Directive 2018/1972: Establishing the European Electronic Communications Code, European Parliament and the Council of the European Union*, December 11, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L1972.

As a lifeline sector, communications networks are tightly integrated with other critical infrastructure sectors, with many sectors having significant dependence on communications. See chapter 1 for an overview of lifeline sectors and their relationship to other critical infrastructure sectors. At the same time, the communications sector is highly dependent on power, particularly reliable commercial electricity. During disasters and disruptions to the electric grid, or in places where reliable power is not available, communications systems rely upon on-site generation, making access to fuel (gasoline, diesel, and sometimes propane) for generators vital to communications resilience. This local generation makes the security of generators and fuel tanks essential and places a premium on maintaining access to remote facilities to refuel and repair generators.

Communications is also highly integrated with the information technology (IT) sector, which provides the hardware and software necessary to run modern communications networks. Communications providers, IT companies, and even content producers both partner and compete to offer massive data-handling capabilities through cloud services, edge-computing capabilities, and content-delivery networks. Communications providers work closely with hardware and software suppliers to manage supply-chain security and resilience.

Communications Industry Segments

Within the broad context of the communications sector, discrete segments or subsectors within communications have unique elements while still sharing common characteristics outlined above. In many cases, private-sector companies focus on a particular subsector, though communications convergence means that the subsectors are often integrated into broader multi-segment systems. Figure 9-1 illustrates the communications sector and role the various industry segments play within this architecture.[2]

2. US Department of Homeland Security (DHS), *2015 Communications Sector Specific Plan* (Washington, DC: DHS, 2015), 5, https://www.cisa.gov/sites/default/files/publications/nipp-ssp-communications-2015-508.pdf.

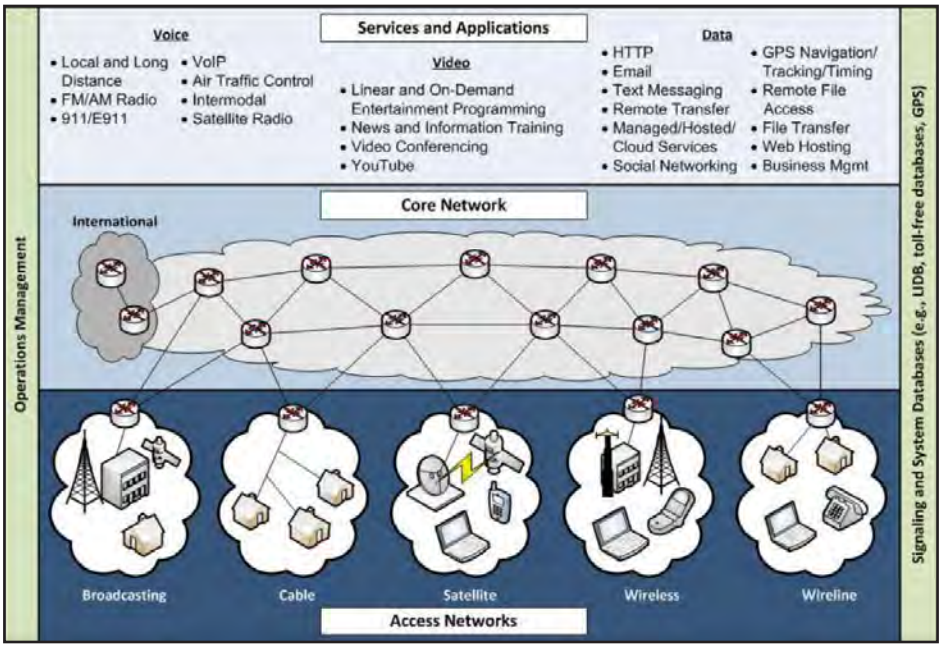


Figure 9-1. Communications sector architecture model
(Diagram by Department of Homeland Security)

The wireline industry segment consists of circuit- and packet-switched networks via copper, fiber, and coaxial transport media. The wireline segment includes private-enterprise data and telephony networks, the core backbone of the Internet (including undersea cable networks), and the public-switched telephone network. Wireline backbone networks have largely completed a transition to fiber-based, packet-switched core networks while using a mix of fiber, coaxial, and twisted pair copper for last-mile connectivity. This modern fiber-optic infrastructure is essential to support the very high bandwidth necessary for the range of voice, video, and data connectivity demanded by government, enterprise, and individual customers.

The cellular communications networks of the wireless segment have undergone significant evolution through 2G, 3G, 4G, and now the implementation of 5G networks. Similar to the wireline segment, most cellular communication—even basic-voice services—is now packet-based Internet protocol traffic. 4G networks revolutionized how cellular networks deliver data services to individual customers. 5G networks will enhance that connectivity, and, perhaps more importantly, will expand that connectivity to include an expected massive proliferation of IoT devices. In addition to cellular communications, the wireless segment also includes wireless hot spots, personal communication services, point-to-point microwave,

high-frequency radio, unlicensed wireless, and other commercial and private radio services.

The satellite segment is also undergoing transformational change. The traditional satellite market remains important, with large, expensive platforms providing a robust suite of data, voice, video, and sensing services, but a combination of lower-cost launch vehicles and highly miniaturized platforms (sometimes known as "CubeSats") has led to a burst of innovation in small satellite network technologies and applications. Space-based position, navigation, and timing (PNT) services from the US Global Positioning System (GPS) and the European Galileo constellations—as well as Russian GLONAST and Chinese BeiDou systems—have become thoroughly integrated into a wide variety of critical infrastructure sector operations.

Today's cable industry provides much more than simple television, offering high-speed broadband and digital telephony over hybrid fiber–coaxial networks. In addition to consumer services, cable providers may also supply enterprise data networks and backhaul services for cellular providers.

The broadcast segment includes free and subscription-based over-the-air radio and TV stations that offer analog and digital audiovisual programming and data services. The transition from analog to digital services in TV and radio allows transmitting of more data using less spectrum, potentially freeing up spectrum for other uses. One such example is the US effort to "re-pack" ultrahigh frequency TV stations into a smaller frequency band to reallocate spectrum for other uses. Radio and TV stations also stream broadcast and additional programming content over the Internet. From an emergency preparedness perspective, it should be noted that traditional broadcast radio is perhaps the most resilient means of communicating with large populations during emergencies because radios are inexpensive and do not require much power and transmitter sites have relatively large broadcast footprints given sufficient transmitter power.

Threats to Communications

Natural Disasters

Natural disasters have the potential to disrupt severely or destroy communications networks either through direct damage to facilities, towers, and cabling or by disrupting the ability to supply power. In the United States, hurricanes and related weather events cause significant local and sometimes regional impacts to communications systems with high winds that damage

aerial cables and power lines, while wind, flood, and storm surge from the most powerful storms devastate core network infrastructure. In some regions, earthquakes threaten communications systems through damage to physical facilities or destruction of conduits underground or along bridges and other rights-of-way while also potentially impacting undersea cables and landing stations. Regardless of the type of disaster and degree of initial damage, communications networks can continue to degrade in the days following a disaster because debris clearance operations often further damage aerial and underground fiber-optic lines while the inability to access facilities (due to physical obstacles or civil curfew) limits the ability to refuel on-site generators and maintain essential facility equipment. Space weather can also significantly degrade communications systems. During periods of intense space weather, sensitive electronics are more likely to be hit and damaged by charged particles. Ions striking satellites can overwhelm sensors, damage solar cells, and degrade wiring and other equipment. Additionally, strong geomagnetic storms can disrupt radio frequency signals such as GPS, satellite TV, high-frequency radio, and even AM broadcast frequencies. At extreme levels, space weather has the potential to induce current flows through long metallic objects, including railroad systems, pipelines, and copper cabling.

Physical Attacks

Like any other critical infrastructure, communications facilities and cable conduits are vulnerable to physical attacks. Remote infrastructure, such as communications towers, mountaintop relay stations, and long-haul fiber-optic runs, can be difficult to protect. Larger facilities, such as central offices and other colocation facilities where network operators interconnect, are generally more protected by physical measures, surveillance, redundant systems, and other resilience measures, but could still experience significant impacts if damaged or destroyed. Satellite gateways and ground stations may also be vulnerable to attack.

Radio frequency–based communications systems—including cellular networks, space-based PNT services, satellite communications, and microwave links—are also vulnerable to jamming, spoofing, and interception. Of these, jamming requires the least technical sophistication as it generally consists of simply overpowering the target signal with a more powerful noise generator. Spoofing is similar to jamming, except the malicious signal mimics the characteristics of the target signal in an attempt to inject false data into the communication system (for example, military capabilities that cause PNT receivers to generate inaccurate location or timing solutions).

Interception does not seek to prevent communication, but instead targets the confidentiality of information passed along the networks.

Submarine cables have unique considerations. They are at the greatest vulnerability in shallow water near shore up through the cable-landing facility. While satellite and deep-water portions of undersea cables are typically beyond the reach of terrorist organizations, advanced nation-state adversaries may have the means to target directly even the most secure critical communications infrastructure—not only deep undersea cables using deep-sea submersibles, but also space-based assets with anti-satellite weaponry.

In addition to intentional attacks, accidental damage can also impact communications networks. Underground communications conduits can be damaged or destroyed by excavation or drilling if cable location-finding and safety protocols are not carefully followed. Post-natural disaster debris clearance and reconstruction in particular can put communications circuits at risk due to the time pressure for restoration and more chaotic nature of operations. Undersea cables, particularly in shallower water, have been damaged by fishing and anchoring operations.

Cyberattacks

Communications systems are also vulnerable to a range of cyberattacks. See chapters 3–5 for more detail regarding cyber threats and actors. Communications is unique in that it can be both a target of cyberattack but also a vector through which other critical infrastructure or systems are attacked. Distributed denial of service attacks attempt to make virtual resources unavailable or compromise defensive measures by flooding network resources with bogus information that consumes bandwidth and computing resources. Some cyberattacks seek to compromise online information systems to target the integrity or confidentiality of data. In some cases, attackers seek to take over control of systems to commandeer or control the operation of critical infrastructure or other vital systems. Ransomware attacks seek to profit by sequestering vital records and systems from their users, then demanding payment to release the systems and data.

Case Studies

To represent the breadth of risks to communications, this section presents five case studies covering different aspects of communications threat and response scenarios across a range of cyber and physical incidents. Each case study is meant to introduce or reinforce essential concepts

for the security and resilience of communications systems, particularly against threats from terrorist actors. While the individual case studies are at a high level, in most cases the references provided will allow a much deeper review and consideration of the specific events. The various scenarios took place worldwide, including in the United States and Europe, within the last 15 years.

Physical Attack: Bombing of a Central Office, Nashville, United States

At 6:30 a.m. on December 25, 2020, a large vehicle-borne improvised explosive device exploded outside an AT&T central office in Nashville, Tennessee. The building survived intact, but physical damage to the lower floors, along with fire and flooding, left the building without power for several days due to damage to commercial power connections to the facility, the blast's effect on the on-site generators, and concerns about internal power distribution systems. An investigation by the Federal Bureau of Investigation (FBI) concluded the attacker had acted alone and was not part of a larger organization, and though the FBI noted the attacker believed in a range of eccentric conspiracy theories, they did not find an ideological motivation for targeting this facility.[3] Ultimately, the attacker's primary motivation appears to have been to end his own life in a high-impact event, while still limiting the risk of injury to others.

Work on restoring communications began almost immediately, with AT&T moving personnel and equipment from its Network Disaster Recovery teams to Nashville while simultaneously working to reroute nonterminating traffic around the facility.[4] Initial onsite recovery efforts were hindered by difficulty accessing the building due to a combination of factors, including safety concerns related to the building's structural integrity and restrictions related to the facility's status as an active crime scene. AT&T provided regular updates on the recovery of the facility on a dedicated public web page and to government and industry counterparts through established information-sharing protocols.[5] By the morning of December 26, AT&T had deployed portable cellular equipment with satellite backhaul to restore connectivity in the region, ultimately ramping up to over 25 portable cellular sites at the height of their response.

3. "FBI Releases Report on Nashville Bombing," Federal Bureau of Investigation Memphis Field Office (website), March 15, 2021, https://www.fbi.gov/contact-us/field-offices/memphis/news/press-releases/fbi-releases-report-on-nashville-bombing.

4. "National Disaster Recovery," AT&T (website), n.d., accessed September 28, 2021, https://www.corp.att.com/ndr/.

5. "Nashville Recovery Efforts," AT&T (website), n.d., March 31, 2021, https://about.att.com/pages/disaster_relief/nashville.html.

AT&T brought in large-scale mobile generators to restore power to the facility, but due to the impacts of the blast, the company had to drill new access holes to route power safely into the building. AT&T restored power to at least four floors of the building by the morning of December 27. By the morning of December 28, just three days following the blast, AT&T stated it had restored the majority of services.

While the attacker may or may not have intentionally selected his location for its proximity to a communications facility, he could scarcely have executed a more targeted attack on a critical communications node. The impacts from the attack stretched across multiple states and affected thousands of customers, including wireless and wireline services to consumers, enterprise customers, and other communications carriers. The Tennessee Emergency Communications Board noted that 66 public safety answering points—call centers where emergency "911" calls for first responder help are answered—were affected.[6] Since the facility hosted colocation network-to-network interfaces, multiple other carriers (wireline, cable, and wireless) had local and regional impacts lasting several days.

There are several key observations and lessons to learn from the Nashville incident and response:

- Coordination between infrastructure operators and first responders is essential, particularly with respect to access to facilities. Where possible, procedures should be worked out in advance regarding access and credentialing.
- Given the interconnected nature of today's communications networks, carrier-to-carrier coordination is critical. In the United States, all major carriers, along with many trade associations and smaller operators, are part of the Communications Information Sharing and Analysis Center (Comm-ISAC), which is partnered with the US Department of Homeland Security's (DHS) National Coordinating Center for Communications. The National Coordinating Center and Comm-ISAC immediately began sharing information on the incident and held the first of several coordination teleconferences on December 26. See chapter 11

6. Donny Jackson, "Tennessee Board Reports Nashville Bombing Impact on 911, Future Plans," IWCE's Urgent Communications (website), May 21, 2021, https://urgentcomm.com/2021/05/21/tennessee-board-reports-nashville-bombing-impact-on-911-future-plans/.

for more information the important role agencies like Comm-ISAC play in sharing information and intelligence.

- Physical hardening of infrastructure works. While buildings across the street from the central office were completely destroyed, the very solidly constructed central office suffered little or no structural damage, allowing very rapid restoration once the building was dewatered and power was rerouted to the still intact floors and equipment.
- The resilience of backup systems, including redundant "last mile" connections, is essential for mission-critical communications requirements
- The ability to reroute rapidly and efficiently, often enabled through software-defined networking, is critical to keeping local disruptions from impacting regional or national communications traffic.

Physical Accident and Attack: Egyptian Undersea Cable Outages (2008 and 2013)

When critical infrastructure extends over long distances—whether power lines, rail links, highways, pipelines, or fiber-optic cables—it can be difficult to protect, as there is typically no way to establish a defensive perimeter around the asset. While both terrestrial and undersea fiber-optic cables can be targets, undersea cables have higher risks because there are fewer alternative pathways, they are more difficult to repair, and their capacity tends to be highly concentrated on relatively few paths. One critical chokepoint for undersea cabling runs along the eastern Mediterranean Sea into Egypt and the Red Sea as figure 9-2 depicts.[7] While accidental damage to undersea cables is commonplace, two noteworthy incidents happened in the vicinity of this chokepoint in 2008 and 2013.

7. "Submarine Cable Map," TeleGeography (website), n.d., accessed September 28, 2021, submarinecablemap.com.

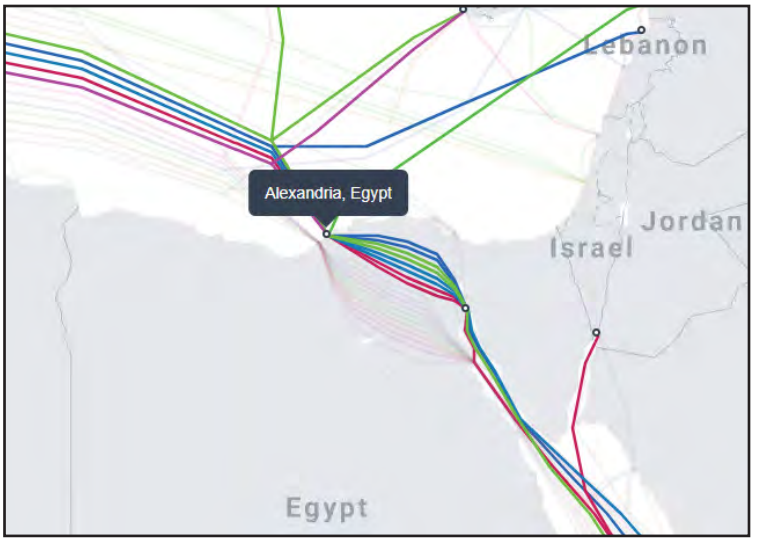


Figure 9-2. Submarine cable map of the eastern Mediterranean Sea
(Map by TeleGeography)

In January and February 2008, a series of apparently unrelated accidents in the eastern Mediterranean Sea and the Middle East caused significant impacts to global information flows. The most consequential of the cluster of cuts took place just off the coast of Alexandria, Egypt, on January 30, 2008, severing two major fiber-optic lines—the SEA-ME-WE 4 and the FLAG Europe-Asia cables—which together represented about 75 percent of the connectivity between the Middle East and South Asia.[8] This event, later attributed to a ship anchoring evolution damaging both cables, degraded Internet connectivity from Egypt (by 70 percent) to India (by 60 percent) with disruptions in the United Arab Emirates (UAE), Kuwait, and Saudi Arabia as well.[9] The effects of this disruption were magnified by several additional cuts to the FALCON cable near UAE, which occurred both before (on January 23) and after (on February 1) the Alexandria incident.[10] On December 5, 2008, another major cut hit SEA-ME-WE 4 and the FLAG Europe-Asia cables, but this time a third cable, the SEA-ME-WE 3, was also severed. Since SEA-ME-WEA 3 had been used to carry rerouted

8. Bobbie Johnson, "How One Clumsy Ship Cut Off the Web for 75 Million People," *Guardian* (website), February 1, 2008, https://www.theguardian.com/business/2008/feb/01/internationalpersonalfinance business.internet.

9. "Severed Cables Disrupt Internet," *BBC News* (website), January 31, 2008, http://news.bbc.co.uk/2/hi/technology/7218008.stm.

10. Rene Wilhelm and Chris Buckridge, "Mediterranean Fibre Cable Cut – a RIPE NCC Analysis," RIPE Network Coordination Centre (website), January 30, 2019, https://www.ripe.net/analyse/archived-projects/mediterranean-fibre-cable-cut.

traffic during the January cuts, the December event was more impactful, affecting the Maldives (100 percent outage), India (82 percent outage), Qatar, Djibouti, and the UAE (70 percent outage each) and with around 50 percent disruption in Saudi Arabia, Egypt, and Pakistan.[11]

At the time, many conspiracy theories flourished in the affected countries and online. While submarine cable cuts from a range of factors happen nearly 200 times per year, the confluence of impacts to cables serving the same regions simultaneously fueled rampant speculation until discarded ship anchors were found near the sites of several of the cuts.[12] Although a direct link had not been demonstrated, it seems as if bad actors were paying close attention to the events of 2008. In 2013, multiple cables in the same area off the Egyptian coast again suffered multiple breaks; this time, however, the Egyptian Coast Guard apprehended three divers trying to cut the SEA-ME-WE-4 cable a few hundred yards off the coast of Alexandria.[13]

A 2017 publication sponsored by the US Office of the Director of National Intelligence and the DHS Public-Private Analytic Exchange Program highlighted a range of threats to submarine cable infrastructure. The report assessed the risk to undersea cables at various points along a given cable route, covering overland last-mile, near-shore, offshore, continental-shelf, and deep-sea sections. Of particular interest in the counterterrorism context is the threat assessed to "vandals, activists and terrorists," which was considered highest in the overland and near shore segments. The report specifically noted that the "concentration of cable landing sites in very few physical locations and the relative ease in finding documented cable routes and cable termination points could facilitate the targeting of the submarine cable network by bad actors."[14] The report also noted that terrestrial portions of long-haul cable networks are also subject to attack.

11. Kim Zetter, "Undersea Cables Cut; 14 Countries Lose Web—Updated," Wired (website), December 19, 2008, https://www.wired.com/2008/12/mediterranean-c/.

12. James Griffiths, "The Global Internet Is Powered by Vast Undersea Cables. But They're Vulnerable," *CNN* (website), July 26, 2019, https://www.cnn.com/2019/07/25/asia/internet-undersea-cables-intl-hnk/index.html.

13. Paul Saffo, "Disrupting Undersea Cables: Cyberspace's Hidden Vulnerability," *New Atlanticist* (blog), April 4, 2013, https://www.atlanticcouncil.org/blogs/new-atlanticist/disrupting-undersea-cables-cyberspaces-hidden-vulnerability/.

14. James Dean et al., *Threats to Undersea Cable Communications* (Washington, DC: DHS, 2017), 8, https://www.dni.gov/files/PE/Documents/1---2017-AEP-Threats-to-Undersea-Cable-Communications.pdf.

There are several options available to mitigate risks to communications cables, particularly from potential targeting by terrorist groups. Some key lessons-learned and risk mitigation options include:

- In close coordination with the private-sector operators of cables, ensure physical and logical route diversity for all critical communications paths and operational preparations to reroute traffic rapidly. In the case of submarine cables, this should also include avoiding, where possible, chokepoints and concentrations of redundant cables, such as near the landing sites at Alexandria, Egypt.
- Maintain sufficient capability to repair critical cables without undue delay.
- Consider options to harden near-shore and onshore submarine cable facilities. Examples include:
  - Trench and bury near-shore segments to help alleviate risks from accidental fishing or anchoring damage, as well as attempts at sabotage by divers or dredgers in shallow water.
  - Harden cable landing stations.
  - Establish proactive coordination among national and local law enforcement, the Coast Guard, the Navy, telecommunications operators, and landing station operators for security and incident response.
  - Minimize publicly available information that provides specific locations of near-shore cables and cable landing stations.
  - Harden related physical infrastructure as much as practicable, including by encasing fiber in concrete, welding manhole covers shut, and securing other potentially exposed components, such as wiring closets, access panels, and elevator shafts, in shared-use facilities.

Natural Disaster: 2017 US Hurricane Season

The 2017 hurricane season was uniquely devastating across the southeastern United States and the Caribbean islands. In fact, the devastation was so significant, that the US Federal Communications Commission took the

extraordinary action of issuing a specific report detailing the many impacts to communications infrastructure. The report focused on the damage from three primary storms: Hurricane Harvey, which hit Houston, Texas, in August 2017; Hurricane Irma, which hit the US Virgin Islands (USVI), Puerto Rico, and Florida in early September 2017; and Hurricane Maria, which decimated the USVI and Puerto Rico in mid-September 2017. Since the most damaging and longest-lasting effects from the three storms were from Hurricane Maria in Puerto Rico and the USVI, this case study will focus on the impacts from Maria.

Hurricane Maria almost completely destroyed the communications infrastructure in both the USVI and Puerto Rico. At its worst, Hurricane Maria took out over 75 percent of cellular service in the USVI and over 95 percent of cellular service in Puerto Rico, where 48 of 78 municipios (county equivalent) had 100 percent loss of cell service.[15] Figure 9-3 illustrates this devastating loss of cellular service in Puerto Rico the day after Hurricane Maria struck.[16] Further, unlike previous storms, cellular service did not bounce back quickly after Maria; in fact, some cellular sites were still out of service six months later. On the USVI, the primary public-safety answering point was out of service for at least 10 days. Radio and TV broadcasters were similarly impacted, with large numbers offline for extended periods of time.

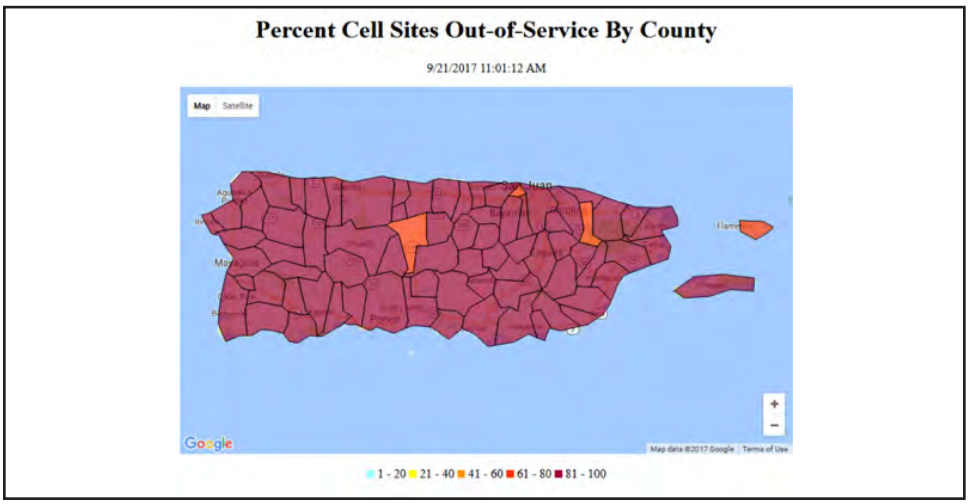


Figure 9-3. Cellular service in Puerto Rico the day after Hurricane Maria

15. Public Safety and Homeland Security Bureau, *2017 Atlantic Hurricane Season Impact on Communications: Report and Recommendations* (Washington, DC: Federal Communications Commission, August 2018), 15, https://docs.fcc.gov/public/attachments/DOC-353805A1.pdf.

16. *2017 Atlantic Hurricane Season*, 16.

(Map by Federal Communications Commission)

While the impacts from a natural disaster may seem unrelated to counterterrorism planning, there are several key lessons learned from these severe weather events:

- Communications resilience is tightly linked to the availability of reliable electric power. By far the primary cause of cellular site failures was lack of commercial power combined with lack of generator capability, generator failure, or lack of fuel.
- Widespread natural disasters will impact communications networks to some degree, and these communications impacts will adversely affect situational awareness and command-and-control necessary to develop and execute response courses of action.
- Communications failures—such as cellular connectivity, the ability to call for first responders, or lack of broadcast capability—have a significant impact on community resilience. The systemic stress to infrastructure systems, including communications, from natural disasters leaves communities more vulnerable to effects from a subsequent attack of opportunity by terrorists, while the impacts of such an attack could be amplified.

Cyberattack on Communication Systems: TV5 Monde

On April 8, 2015, a group claiming to be the "Cyber Caliphate" attacked the French television network TV5 Monde. The groundwork for the sophisticated attack actually began months earlier, with initial network penetration happening as early as January 23, 2015. The attackers conducted systemic reconnaissance of the TV5 Monde network to understand how the network developed, staged, and broadcast its content. Systematically burrowing deeply into TV5 Monde's network, the attackers built malicious software that corrupted the Internet-connected hardware that controlled the station's operations, even compromising the station's ability to encode content for broadcast.[17] When the main attack was launched, all 12 of the network's broadcast channels were initially taken off the air,

17. Gordon Corera, "How France's TV5 Was Almost Destroyed by 'Russian hackers,'" *BBC News* (website), October 10, 2016, https://www.bbc.com/news/technology-37590375.

with some channels not restored for 18 hours. Nearly simultaneously, the network's website and social media accounts were also compromised and used to post propaganda supporting the Islamic Caliphate.

While contemporaneous accounts of the event initially expressed alarm at an apparently and suddenly more capable and sophisticated cyber adversary in the Islamic Caliphate, signs began to emerge that the attack may have been a "false flag" operation conducted by another actor trying to shift blame for the attack. The complex attack used multiple points of ingress in the attempt to compromise TV5 Monde's network, even including the station's Internet-connected security cameras. The computer forensics company, FireEye, finally attributed the attack to the Russian advanced persistent threat (APT) group known as Fancy Bear or APT 28, suggesting the false flag effort was "likely a Russian information operation" meant to "capitalize on Western fears over Islamic extremism" and "draw the West's attention away from Russia's ongoing role in the Ukraine crisis and towards the threat of terrorism in the Middle East."[18] FireEye, along with government investigations and other cybersecurity firms, noted a range of evidence pointing to Russian involvement, including Internet Protocol (IP) blocks adjacent to other known APT 28 activity, code compilations consistent with operations in Moscow and St. Petersburg, and a code base written with Cyrillic keyboards.[19]

The net result of the incident was that the attacker took 12 TV5 Monde channels off the air and caused a direct estimated financial impact of over 4.5 million euros. Fortunately, attackers either could not or chose not to commandeer the broadcasts to air content of their choosing. The attackers did, however, commandeer related TV5 Monde social media accounts (YouTube, Twitter, and Facebook) to launch propaganda campaigns. Some observations and lessons learned from the APT 28 attack include:

- The importance of multilayered defense. According to multiple sources, TV5 Monde did not have good segmentation of their internal network. Their information technology and operations technology systems were interconnected, allowing a breach on the business side of the network to impact to operational systems.

18. FireEye, *Cyber Threats to the Entertainment and Media Industries* (Milpitas, CA: FireEye, 2016), 1–2, https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/ib-entertainment.pdf.

19. "Russian Hackers Likely behind 'IS Group Cyber Attack' on French TV Network," *France 24* (website), October 6, 2015, https://www.france24.com/en/20150610-france-cyberattack-tv5-television-network-russia-hackers.

- In cyber events, attribution is difficult and, even in the best circumstances, leaves a degree of uncertainty. Further, even where attribution is possible, publicly acknowledging the attribution may reveal the sensitive sources and methods through which attribution was achieved. These challenges with attribution complicate deterrence since placing consequences on an attacker requires some degree of certainty that the actual attacker is bearing the cost of the retribution. See the Ukraine case study in chapter 5 for discussion on the difficulty of attribution and response options to cyberattacks.
- While communications systems are often involved in cyberattacks designed to reach other targets, communications are sometimes targeted directly. As both nation-states and non-state actors continue to view information and disinformation campaigns as a means to achieve their ends, assuming direct control of the means of communication will continue to be an attractive target.
- The ability to attack broadcast networks, in conjunction with other actions in a complex attack, could greatly hinder the ability to communicate to the public, thereby causing significantly more impact than a simple kinetic attack alone.

Distributed Denial of Service (DDoS): Mirai Botnet

In early August 2016, a new malware variant, known as Mirai, began to circulate. This self-replicating malware specifically targeted IoT devices such as home routers, security cameras, printers, and network-attached storage devices. Once installed on a new device, the malware would seek other devices to infect by randomly scanning the Internet for targets and then attempting to gain access by trying 64 commonly used login and password combinations. Many IoT devices at the time used default passwords that were rarely changed by users, so the malware was able to spread remarkably rapidly. The cybersecurity company, Cloudflare, assessed that during its first few days, the malware was doubling the number of devices captured every 76 minutes.[20] These captured devices (now known as bots) would form a sort of network (a botnet) when each device would contact command-and-control (C2) servers to receive instructions on what to do next. In the case of Mirai, the C2 servers could provide instructions to launch DDoS attacks to flood

20. Elie Bursztein, "Inside the Infamous Mirai IoT Botnet: A Retrospective Analysis," *Cloudflare* (blog), December 14, 2017, https://blog.cloudflare.com/inside-mirai-the-infamous-iot-botnet-a-retrospective-analysis/.

target computers and networks with bogus traffic to overwhelm the servers' ability to respond, thus making a site or service unavailable for legitimate use.

The first two victims of the Mirai botnet were the European web-hosting company OVH on September 18, 2016, and the servers hosting the online presence for noted cybersecurity researcher Brian Krebs on September 21. At the time, the attacks were some of the largest DDoS attacks ever observed. Later that month, Mirai's creator posted its source code online, leading to a massive proliferation of botnets built on the Mirai model, and it did not take long for these new botnets to become active. On October 21, Mirai was used to target Dyn, a domain name system (DNS) resolver. DNS resolvers such as Dyn translate the text-based Internet addresses into the numerical IP addresses to enable the routing of Internet traffic. Since Dyn provided this service for so many customers, the attack on their DNS service made large swaths of the Internet unavailable for a short period until the attack was mitigated. In the months that followed, other attacker groups using Mirai targeted a wide range of victims across the globe. Mirai variants remain active to this day; in the second quarter of 2021 alone, global communications company, Lumen, through its Black Lotus Labs threat research arm, identified 349 unique Mirai C2 infrastructures, with each Mirai C2 "family" attacking an average of more than 15,000 victims over a typical one-month lifespan.[21]

For several reasons, the Mirai botnet is an important example to understand when considering the terrorist cyber threat to communications systems:

- Once "in the wild," the exploit was rapidly weaponized by a large number of groups, at least some of whom sold DDoS services for hire, putting this kind of cyberattack capability in the hands of any individual or group willing to pay for it.
- The malware successfully promulgates itself in environments where basic cybersecurity measures are lacking (for example, where Internet-facing devices with default or common passwords are left exposed).
- The attack itself was not technically sophisticated. The original code was written by an undergraduate student and two friends. An FBI agent investigating the case said, "These kids are super smart, but they didn't do anything high

21. "Lumen Quarterly DDoS Report Q2 2022," Lumen Technologies (website), https://assets.lumen.com/is/content/Lumen/lumen-quarterly-ddos-report.

level—they just had a good idea."[22] While state-sponsored cyber actors sometimes wield remarkably sophisticated capabilities, cyberspace is a battlefield where smaller players can create, borrow, or buy relatively simple capabilities that can have outsize impacts on CI security and resilience.

Conclusion

As described in this chapter, communications systems are a foundational component of critical infrastructure that are essential to a functioning economy and state capacity to conduct crisis response and military action. Based on recent incidents affecting communications systems, such as those discussed in the case studies, there are several important actions governments and industry can take to promote communications sector security and resilience.

Blue-sky Coordination and Relationship Building

The security and resilience of the communications sector is immeasurably improved by a vibrant and strong public-private partnership among communications providers, national-security professionals, and public-safety officials. These relationships must be established and nurtured well ahead of an impending disaster or actual attack. Routine and regular interaction is necessary to build personal familiarity and establish trust. This interaction also ensures that coordination processes and procedures are up to date, to include even things as basic as having contact information, distribution lists, and virtual platforms that are available and functional. In the United States, the DHS National Coordinating Center for Communications hosts a standing weekly call with major federal departments, communications regulators, and dozens of communications company representatives; during the call, developing cyber and physical incidents are discussed, questions asked and answered, and a common lexicon and set of understanding is established. Partnerships developed through mechanisms such as this form a foundation upon which the range of other problems and opportunities can be most easily addressed, such as:

22. Garrett M. Graff, "How a Dorm Room Minecraft Scam Brought down the Internet," Wired (website), December 13, 2017, https://www.wired.com/story/mirai-botnet-minecraft-scam-brought-down-the-internet/.

- Establishing basic coordination mechanisms via teleconferences, video conferences, in-person meetings, e-mail distributions and online portals, which are used to share information and coordinate response during steady-state operations and times of crisis.
- Pre-establishing disaster-reporting processes, so all key stakeholders understand the expected cadence of reporting and have access to reporting templates and application programming interfaces to build internal reporting processes.
- Creating a centralized coordination mechanism to minimize multiple reporting and request paths between government and industry.
- Exercising jointly developed mechanisms before using them in crisis.
- Encouraging industry-to-industry coordination and cooperation that can be leveraged during incident response.

Identification of Risks and Appropriate Mitigation Strategies

As with all critical infrastructure risk-management activities, promoting the security and resilience of communications is most effective when the owner-operators of communications networks are engaged early in the process. Owner-operators, armed with the best available information about threats to their infrastructure, are best positioned to make the most efficient investments in security and resilience initiatives to address those threats. Similarly, those reliant on communications systems benefit from a clear analysis of the risks to those systems on which they rely.

A review of recent events, including those presented in the case studies here, show several areas where investments in resilience have been particularly effective:

- Physical and virtual path diversity. Critical communications should have diverse physical and logical paths to facilitate greater resilience. Today's meshed networks provide excellent route diversity over long-haul core networks, with software-defined networking enabling flexibility of these routes to allow for rapid traffic engineering away from disrupted networks. Such dynamic rerouting is only possible, however, where physical infrastructure allows and so presents a greater challenge in less dense areas, particularly

over the "last mile" legs connecting end users. Those who rely on connectivity for mission-critical military, public safety, or business requirements should work closely with network providers to understand and accommodate physical and logical route diversity.

- Cybersecurity baseline practices. Ensuring a foundational level of cybersecurity best practices—such as routine patching, firewall controls, intrusion detection, and multifactor authentication—will help to prevent or at least mitigate attacks by cyber criminals and low-level non-state actor groups. High-level state-sponsored attackers and increasingly capable attackers-for-hire may be able to breach even the best cyber defenses, making impact mitigation strategies—including zero trust architectures, network segmentation, least privilege implementation, and data encryption—equally important. See chapter 14 for an overview of recommended cybersecurity best practices and tools.
- Cross-sector relationships among communications, electricity, and information technology. Most of the significant events affecting communications over the past 20 years involved concurrent issues with either electric power or information technology. It is critical that the communications sector routinely coordinates and addresses systemic risks at the intersection of these three sectors and that government security and response planners factor this need for coordination into operational and information-sharing plans.

Communications Sector Resilience Enablers

Commercial communications providers actively manage network events on a daily basis and are the experts on protecting, maintaining, and restoring their networks. There are, however, key areas where government and industry can partner to foster improved sector resilience.

In addition to the partnership mechanisms described above, the private sector can benefit significantly from regular information sharing from governments about known threats from adversaries before incidents occur. Understanding likely adversary tactics, techniques, and procedures will help industry develop and implement tailored countermeasures and risk mitigations. Where specific pre-attack tactics, techniques, and procedures or other cyber indicators of compromise are known, they should be shared with industry

so owners and operators can be on the lookout for signs of an impending attack. Government and industry can also benefit from rapid sharing of prospective defensive measures, so they can evaluate effectiveness within their industry or facility and implement quickly where it makes sense to do so. See the discussion of information-sharing benefits and best practices in chapter 11.

During incidents, experience has shown that industry response is improved by close coordination among relevant government and industry partners. Key elements of this cooperation are:

- Access. Often simply getting personnel and vehicles to the network facilities is a challenge following an attack or disaster or during periods of heightened tension. It is critical that communications providers and government officials work together to enable entry to restricted areas, coordinate road clearance and debris removal, and clear access routes to remote sites as soon as possible.
- Energy. Prompt and/or prioritized restoration of electric power will significantly increase the communications sector's ability to recover quickly, so coordination on power restoration priorities and estimated restoration times is essential. Where electric power is unable to be restored quickly, access to fuels (diesel, gasoline, and sometimes propane) for on-site generation becomes critical.
- Security. The security of communications facilities, including temporary assets such as portable generators, mobile cellular equipment, work crews, and fuel depots, is vital.
- Integrity of communications conduit and cabling. During post-incident operations like debris removal, past incidents demonstrate that fiber-optic and other cables are often damaged or destroyed unintentionally. Close coordination and clear public messaging urging care and caution working in and around fiber conduit and overhead cabling can help mitigate some of this risk.
- Timely information updates. As situations evolve, when government has actionable information—intelligence or identified protective measures, for example—it must share this information with industry quickly so it can be put to use.

As a lifeline sector, communications enables much of the modern world's core functions, including military command and control, government operations, emergency response coordination, economic productivity, and societal engagement. For precisely these reasons, communications are an attractive target for terrorist groups and other adversaries. While communications companies expend great efforts in building secure and resilient networks, government and industry can coordinate efforts to establish solid operational relationships, identify and mitigate critical risks, and develop viable plans of action for disaster response scenarios.

— 10 —

Comparing Policy Frameworks: CISR in the United States and the European Union

Ronald Bearse and Alessandro Lazari

For over a quarter century the United States and the European Union have been diligently planning and implementing policies and procedures to protect the critical infrastructure sectors that are vital to the prosperity and security the majority of their citizens enjoy. Given the evolving nature of threats against critical infrastructure, recent US and EU efforts have focused on enhancing collective critical infrastructure security and resilience (CISR) posture. The core objective of these CISR initiatives is to strengthen their ability to deter, prevent, reduce the consequences of, respond to, and recover from a broad array of vulnerabilities, hazards, and threats to critical infrastructure. Any such disruptions to or destruction of these critical infrastructure systems and assets can have damaging impacts on individual nations, the transatlantic economy and security environment, and the ability of the North Atlantic Treaty Organization (NATO) to fulfill its core tasks.

The US and EU CISR policies and practices have long been recognized as two of the most advanced frameworks in the world. Since a considerable majority of national CISR policies and plans from around the world reflect or reference to some extent the US or EU models, this chapter describes the key underpinnings and characteristics of each framework. In fact, among NATO's 30 member states, 22 of them—the United States and 21 others who are also members of the EU—are directly impacted by these frameworks, while the remaining

eight member states and numerous partner nations are influenced by them.[1] This chapter will first examine the US framework and then that of the EU, with each section outlining the fundamental elements of the respective models, the reasons why these frameworks came into being and how they were adapted over time, and the various ways in which they are being implemented to strengthen national, regional, and international security, economic prosperity, and public health and safety. The goal of this chapter ultimately is to help Allies and partners better understand these two frameworks and apply their key principles and tenets to enhance the CISR posture in their respective countries.

US CISR Framework

Critical infrastructure first appeared in formal US policy in 1996 when US President Bill Clinton signed Executive Order 13010, which established a national commission to assess the scope and nature of the vulnerabilities of and threats to critical infrastructure facilities and systems. The commission also recommended a comprehensive national policy and implementation strategy for protecting critical infrastructure from physical and cyber threats and for assuring their continued operation.[2] Five years later, the USA Patriot Act of 2001 took the additional step of defining critical infrastructure as those "systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters."[3]

More than a decade later, in 2013, US President Barack Obama established current US national policy on CISR when he signed Presidential Policy Directive 21 (PPD-21) to advance a national unity of effort to strengthen and maintain secure, functioning, and resilient critical infrastructure against all hazards, including physical and cyber threats.[4] PPD-21 formally

1. "Relations with the European Union," North Atlantic Treaty Organization (website), June 21, 2021, https://www.nato.int/cps/en/natohq/topics_49217.htm.

2. "Critical Infrastructure Protection," Exec. Order 13010, July 15, 1996, https://www.presidency.ucsb.edu/documents/executive-order-13010-critical-infrastructure-protection.

3. Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT ACT) Act of 2001, H.R. 3162, (2001), https://www.congress.gov/107/plaws/publ56/PLAW-107publ56.pdf.

4. White House, *Presidential Policy Directive (PPD)-21—Critical Infrastructure Security and Resilience* (Washington, DC: Office of the Press Secretary, February 12, 2013), https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil/.

marked the distinction between critical infrastructure protection (CIP) and making it more secure and resilient (CISR).

Critical infrastructure in the United States consists of distributed networks, numerous types of organizational structures and operating models (to include multinational ownership), interdependencies in the physical and cyber domains, and unique arrangements comprised of authorities, responsibilities, and regulations across all levels of the government and the private sector.[5] Given this complex and diverse nature of critical infrastructure, the policy aim outlined in PPD-21 is ambitious and challenging to accomplish. Therefore, PPD-21 recognizes the need for cooperation between the public and private sectors, coordination among various agencies and levels of government, and integration with the national preparedness system across the spectrum of the prevention, protection, mitigation, response, and recovery domains.[6]

According to the current CISR policy established by PPD-21, the federal government is responsible for strengthening the security and resilience of its own critical infrastructure, ensuring the continuity of national essential functions, and continuing to partner effectively with critical infrastructure owners and operators to enhance their CISR efforts.[7] In addition to its partnership with private-sector owners and operators, the federal government works with state, local, tribal, and territorial entities to manage risks and strengthen the security and resilience of the nation's critical infrastructure against all hazards that could have a debilitating impact on national security, economic stability, or public health and safety. PPD-21 also requires the federal government to engage with international partners to strengthen the security and resilience of domestic critical infrastructure as well as those facilities or assets located outside of national borders upon which the country depends.

What Guides US CISR Policy?

PPD-21 outlines three strategic imperatives that are the foundation for improving national CISR practices and procedures. The first imperative is to enhance functional relationships across the federal government to advance a national unity of effort. Key to this effort is a national plan that identifies roles and responsibilities for sector-specific agencies; other federal departments

5. "Fact Sheet: Executive Order on Cybersecurity/Presidential Policy Directive on Critical Infrastructure Security and Resilience," Department of Homeland Security (DHS) (website), February 13, 2013, https://www.dhs.gov/news/2013/02/13/fact-sheet-executive-order-cybersecurity-presidential-policy-directive-critical.

6. White House, *PPD-21*.

7. White House, *PPD-21*.

and agencies with critical infrastructure roles; state, local, tribal, and territorial entities; and critical infrastructure owners and operators. As the vulnerabilities, risks, and threats to critical infrastructure have evolved in the twenty-first century, CISR policy has also adapted to address these changes, often in the form of innovative programs and initiatives addressing specific infrastructure issues and priorities. PPD-21 outlines the need to establish baseline capabilities that reflect this evolution of knowledge and practice, define relevant federal program functions, and take steps to facilitate collaboration and information exchange between and among the federal agencies.[8] As part of this updated structure, PPD-21 directed the creation of two national centers—for physical and cyber aspects of infrastructure, respectively—under the Department of Homeland Security (DHS) to be focal points for information and situational awareness for critical infrastructure partners. Since its creation in 2018, the Cybersecurity and Infrastructure Security Agency (CISA) fulfills these two functions.

The second imperative, along these lines, is to enable the efficient exchange of information, including intelligence, between all levels of government and critical infrastructure owners and operators to enable situational awareness and multidirectional sharing of threats and vulnerabilities. To enhance multidirectional information exchange between and among the government and private sector, this imperative highlights the importance of identifying (1) requirements for data and information formats and accessibility, (2) system interoperability, and (3) redundant systems and alternate capabilities should there be a disruption in the primary systems.[9]

Building on the first two imperatives, the third strategic imperative calls for the implementation of an integration and analysis function for critical infrastructure that includes operational and strategic analysis of incidents, threats, and emerging risks. Integration and analysis of information resides within CISA's purview, and it includes the capability to collate, assess, and integrate information regarding vulnerabilities and consequences with threat streams and hazard information. While not replicating the analysis conducted in the broader national intelligence community, this function (1) helps prioritize assets and manage risks to critical infrastructure, (2) anticipates cascading impacts due to interdependencies, (3) recommends CISR measures prior to, during, and after an event or incident, and (4) supports incident management and restoration efforts related to critical infrastructure.[10]

8. White House, *PPD-21*.

9. White House, *PPD-21*.

10. White House, *PPD-21*.

This function depends on stakeholders—federal departments and agencies as well as analytic entities at all other levels of government and in the private sector—supplying relevant, timely, and appropriate information to CISA so it can maintain and share near real-time situational awareness with actionable information about imminent threats, significant trends, and incidents that may affect critical infrastructure.[11]

Given these three strategic imperatives and the broader strategic direction outlined in PPD-21, the National Infrastructure Protection Plan (NIPP) 2013 fulfills PPD-21's clear call for an updated national plan. NIPP 2013 presents the vision and mission for CISR policy, for which CISA has primary responsibility among federal agencies. According to the NIPP 2013, the physical and cyber critical infrastructure should "remain secure and resilient, with vulnerabilities reduced, consequences minimized, threats identified and disrupted, and response and recovery hastened."[12] In turn, this vision drives the basic approach to enhance CISR "by managing physical and cyber risks through the collaborative and integrated efforts of the critical infrastructure community."[13]

An important first step is to determine which infrastructure sectors are both critical to maintain continued services or functionality and vulnerable to some type of threat or hazard. The US government designated four sectors—transportation systems, water and wastewater systems, energy, and communications—as lifeline sectors because most other sectors depend on these functions and services to operate. See chapter 1 for its explanation of lifeline sectors.[14] Due to these dependencies and interdependencies between infrastructure elements, the loss of one lifeline function typically has an immediate impact on operations in multiple sectors. See chapter 12 for greater detail on the nature of dependencies, interdependencies, and cascading or escalating effects. Naming and officially recognizing lifeline sectors and identifying existing cross-sector interdependencies facilitates collaboration and information exchange and promotes continuity of operations and services.[15]

11. White House, *PPD-21*.

12. DHS, *NIPP 2013: Partnering for Critical Infrastructure Security and Resilience* (Washington, DC: DHS, 2013), 5, https://www.cisa.gov/sites/default/files/publications/national-infrastructure-protection-plan-2013-508.pdf.

13. DHS, *NIPP 2013*, 5.

14. DHS, *NIPP 2013*, 9.

15. Cybersecurity and Infrastructure Security Agency (CISA), *A Guide to Critical Infrastructure Security and Resilience* (Washington, DC: CISA, 2019), 4, https://www.cisa.gov/sites/default/files/publications/Guide-Critical-Infrastructure-Security-Resilience-110819-508v2.pdf.

In total, US CISR policy currently recognizes 16 sectors as critical infrastructure. In addition to the four lifeline sectors, the additional 12 sectors are: chemical, commercial facilities, dams, defense industrial base, emergency services, energy, financial services, food and agriculture, government facilities, healthcare and public health, information technology, critical manufacturing, and nuclear reactors, materials, and waste (see figure 10-1).[16] Although the number of identified critical infrastructure sectors is currently 16, this number is subject to change regarding the vulnerability of specific sectors and the nature of interdependencies between them. For instance, in 2017, the US government designated election infrastructure as a subsector of the government facilities sector due to the importance of free and fair democratic elections as a foundation of the American way of life.

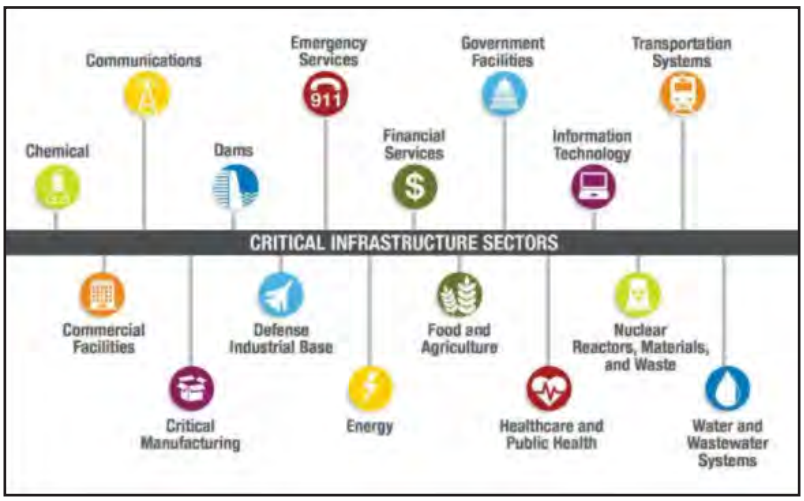


Figure 10-1. 16 critical infrastructure sectors recognized by US policy
(Source: Diagram by DHS)

When it comes to securing these critical infrastructure sectors and making them more resilient, it is essential to understand the nature of the threats and hazards that can influence operations. Some hazards and threats are specific to geographic regions, others affect the entire country, and a few may even have global impacts. By using an all-hazards approach, US CISR policy begins with an appreciation for the spectrum of threats and hazards to critical infrastructure, an analysis of the likelihood of occurrence and their potential impacts, and then directs efforts to focus on and prepare for those that pose the greatest risk. Although not an exhaustive list, the government considers the following hazards and threats when determining appropriate

16. DHS, *Critical Infrastructure Threat Information Sharing Framework* (Washington, DC: DHS, 2016), 6, https://www.cisa.gov/sites/default/files/publications/ci-threat-information-sharing-framework-508.pdf.

CISR policies and practices. See chapters 2–4 for more detail on physical, cyber, and hybrid threats.[17]

- Climatological, hydrological, meteorological, and geophysical events, such as drought, wildfires, floods, tropical cyclones, severe winter storms, earthquakes, tsunamis, or volcanic eruptions
- Pandemics
- Industrial accidents like structural failures and chemical spills
- Unscheduled disruptions due to aging infrastructure or malfunctions
- Criminal incidents and physical or terrorist attacks
- Cyber incidents, including denial-of-service attacks, malware, and phishing
- Attacks that exploit supply-chain vulnerabilities to cause system failure
- Foreign operations to spread misinformation, undermine democratic processes, or make investments that give foreign powers undue influence

Adopting a Sound Risk Management Framework

An essential first step in developing CISR policy is understanding the nature of the risks to critical infrastructure and developing appropriate measures to mitigate or respond to them. The NIPP 2013 defines risk as the potential for an unwanted outcome resulting from an incident, event, or occurrence, as determined by its likelihood of occurrence—given the nature of threats and vulnerabilities—and the consequences that could follow.[18] Risk management is the process of identifying, analyzing, and communicating risk and then accepting, avoiding, transferring, or controlling it at acceptable levels and costs.[19] Risk management enhances US CISR posture in several important ways. First, it brings attention to those threats and hazards that are most likely to cause significant, unwanted outcomes to a specific infrastructure or sector. Second, it informs actions and guides the application of resources to prevent or mitigate the effects of these threats and hazards.

17. CISA, *Critical Infrastructure Security*, 7.
18. DHS, *NIPP 2013*, 7.
19. DHS, *NIPP 2013*, 7.

Third, risk management enables stakeholders to identify and prioritize actions to ensure continuity of essential functions and services and support enhanced response and recovery efforts following incidents. Finally, risk management facilitates decision making and the setting of priorities among all stakeholders.

The current US critical infrastructure risk management framework, described in NIPP 2013, consists of the following core tenets.[20]

- Risk should be identified and managed in a coordinated way within the critical infrastructure community to enable effective resource allocation.
- Partnerships can improve understanding of evolving risk to cyber and physical systems and assets and can offer data and perspectives from various stakeholders.
- Understanding and addressing risks from cross-sector dependencies and interdependencies is essential to enhancing overall CISR posture.
- Gaining knowledge of and reducing infrastructure risk requires information sharing across all levels of the critical infrastructure community.
- A partnership approach, involving public and private stakeholders, recognizes the unique perspectives and comparative advantages of the diverse critical infrastructure community.
- Regional, state, and local partnerships are crucial to developing shared perspectives on existing gaps and actions required for improvement.
- Critical infrastructure transcends national boundaries and thus requires bilateral, regional, and international collaboration, capacity building, mutual assistance, and other cooperative agreements, such as the Canada-US Action Plan for Critical Infrastructure.
- The design phase of critical infrastructure facilities or systems should consider and incorporate measures to enhance security and resilience.

20. DHS, *NIPP 2013*, 13–14.

A New Approach: Managing Cross-sector Risk to Critical Infrastructure

Effective risk management depends on critical infrastructure stakeholders' ability to engage across sectors to facilitate a shared understanding of risks and integrate a wide range of activities to manage them. In April 2019, CISA published a list of 55 national critical functions, which are government and private-sector operations and services so vital to the United States that their disruption, corruption, or dysfunction would have a debilitating effect on physical and economic security and public health or safety.[21] Identifying these critical functions marks a shift from the previous risk-management framework focused primarily on entity-level risk management to more of an enterprise approach that looks instead at critical outcomes.

The national critical-functions framework captures risk and associated dependencies that are cross-sector in nature and may have cascading impacts within and across sectors. Using this functional approach to risk allows the critical infrastructure community to enhance CISR posture in a more strategic way, including more precise targeting and prioritizing of assets and systems based on an understanding of critical functions and key dependencies and interdependencies.

CISA organizes these functions into four domains: (1) technological connections that enable vital communications, (2) distribution methods for the movement of goods, people, and utilities, (3) management processes linked to national security and public safety, and (4) supply chains and services that secure the economy, such as water or housing.[22] The functional framework enables CISA to better identify risks that might otherwise be overlooked, including risks to supply chains, major cybersecurity issues like attacks to steal intellectual property or manipulate industrial-control systems, and the need for cross-industry engagement over complex challenges like vulnerabilities associated with position, navigation, and timing systems.[23]

Who Is Responsible for CISR Efforts?

In the United States, CISR is a collective responsibility that critical infrastructure owners and operators, government entities, and nongovernmental organizations (including industry associations) share.

21. "National Critical Functions," CISA (website), n.d., accessed September 28, 2021, https://www.cisa.gov/national-critical-functions.

22. "National Critical Functions."

23. CISA, *Critical Infrastructure Security*, 8.

Since private-sector companies own and operate most critical infrastructure, public-private partnerships are essential for effective CISR efforts. Although not all-inclusive, this section will introduce some of the most important elements of collective responsibility for CISR efforts and public-private partnerships.

Starting with the government side, sector-specific agencies are those federal entities that have an assigned responsibility for one or several critical infrastructure sectors.[24] For instance, the Department of Health and Human Services is responsible for the healthcare and public-health sector, while the Environmental Protection Agency is responsible for water and wastewater services. Beneath the federal level, governments at the state, local, tribal, and territorial levels play key roles in protecting public safety and welfare, providing essential services, and planning and implementing activities that ensure the security and resilience of critical infrastructure within their respective jurisdictions.[25]

From the nongovernmental perspective, several communities play key roles in enabling CISR efforts.[26] The academic community's contributions include research and development, testing and evaluation of CISR technological advances, and participation in the risk-analysis and management process. Advisory councils, such as the National Infrastructure Advisory Council, provide recommendations, advice, and expertise to government agencies on a host of CISR initiatives and policies.

Finally, critical infrastructure owners and operators from the public and private sectors are essential to a strong CISR posture. Private-sector companies own and operate roughly 85 percent of the physical and cyber infrastructure, while federal, state, or local governments manage the remaining 15 percent.[27] Since private companies own and operate the majority of critical infrastructure, they are uniquely positioned to manage risks to their respective systems and assets and establish effective strategies to make them more secure and resilient.

Perhaps the strongest element of the collective responsibility for CISR practices in the United States is the network of partnerships, beginning

24. DHS, *NIPP 2013*, 43.

25. DHS, *NIPP 2013*, 46–48.

26. DHS, *NIPP 2013*, 48–50.

27. Government Accountability Office (GAO), *Critical Infrastructure Protection* (Washington, DC: GAO, 2006), 1, https://www.gao.gov/assets/gao-07-39.pdf.

with the series of public-private partnership councils.[28] First, sector coordinating councils are self-governed groups of private-sector owners and operators and trade association representatives from the 16 critical infrastructure sectors. Given their collective expertise and management of daily operations for most critical infrastructure facilities, these councils provide the government with key insights and recommendations for planning and implementing sector-specific CISR activities.

Second, government coordinating councils are sector-specific groups that include representatives from all levels of government that serve as the public-sector counterpart for the private-sector councils. Beyond what the sector coordinating councils can do on their own, these government councils enable coordination across jurisdictions and among various government agencies.

Third, cross-sector councils—typically comprised of the primary or vice chairpersons from each of the sector councils—are a venue to address the issues and interdependencies that involve multiple critical infrastructure sectors and share general best practices. Finally, the Regional Consortium Coordinating Council is a national forum that provides the framework and foundation for cross-sector coordination and CISR efforts at the regional level. Given the importance of public-private partnerships, the Critical Infrastructure Partnership Advisory Council (CIPAC) provides the framework to facilitate interaction between these different councils and their various members. CIPAC directly supports NIPP 2013 and PPD-21 by providing a forum in which government agencies and critical infrastructure owners and operators can meet to plan, support, and coordinate CISR efforts.[29] Figure 10-2 on the next page illustrates the relationship between these various councils, federal agencies, and the 16 critical infrastructure sectors.

While CISR programs can be voluntary, regulatory, or some combination of both, voluntary programs are the most common in the United States because they typically work best to promote innovative concepts, especially when vast diversity in a sector or industry limits the ability to apply common standards.[30] Regulatory programs generally are the optimal choice to establish a common standard in the sector or industry, to promote certain industry practices, or to ensure that organizations do not suffer a competitive disadvantage for compliance. The chemical sector, for instance, promotes preparedness

28. "Critical Infrastructure Sector Partnerships" CISA (website), n.d., accessed January 25, 2022, https://www.cisa.gov/critical-infrastructure-sector-partnerships.

29. "Critical Infrastructure Sector Partnerships."

30. DHS, *NIPP 2013*, 10.

through a voluntary framework between industry and government and is partially subject to regulatory programs to ensure compliance to standards.[31]

Figure 10-2. Partnerships: Sector and cross-sector coordination
(Diagram by DHS)

31. "Chemical Sector," CISA (website), n.d., accessed October 4, 2021, https://www.cisa.gov/chemical-sector.

Stakeholders' roles and responsibilities in CISR vary widely and depend on several important factors, such as:

- Whether the critical infrastructure is publicly or privately owned
- Regulations governing operations within a critical infrastructure sector
- Sector-specific threats and hazards
- Whether the sector or region prioritizes actions to protect infrastructure, reduce consequences, or rapidly respond to and recover from adverse events

Industry associations often play a key role in recommending practices, while in other sectors there may be regulations that require owners and operators to take certain actions. Some sectors have statewide or national design standards to protect facilities and assets against damage from events like fires, floods, and earthquakes. Insurance providers may also impose security requirements on their policyholders in some sectors. Depending on the stage of CISR activity—preparation for, prevention against, response to, or recovery from an incident or event—stakeholders have different roles and responsibilities. For example, though first responders, critical infrastructure owners and operators, and regional and federal resources may drive response efforts when an incident occurs, the responsibility for recovery in a voluntary system falls to the owners and operators who know the infrastructure best.

This brief section shows the great value of engagement between private industry and all levels of government and the key role councils play in fostering mutual understanding and trust while promoting information sharing and practical exchanges. In particular, organizations that promote planning, prioritization of resources, exercises, and training contribute to better national preparedness, increasingly secure and resilient infrastructure, and more timely responses, which are ultimately the desired outcomes of CISR policy.

Effective CISR: Built on Collaboration and Information Sharing

As outlined in the NIPP 2013, security is the process of reducing the risk to critical infrastructure from intrusions, attacks, or the effects of natural or man-made disasters by applying physical means or defensive

cyber measures.[32] Similarly, resilience is the ability to prepare for and adapt to changing conditions, including the ability to withstand and recover rapidly from disruptions, deliberate attacks, accidents, or naturally occurring threats or incidents.[33] To truly make critical infrastructure more secure and resilient requires a strong foundation of collaboration and information sharing.

Effective collaboration requires several vital elements: (1) structures and processes to enable participants to communicate freely without releasing proprietary information or providing unfair advantage, (2) a trusted environment where stakeholders share information needed to strengthen security and resilience, and (3) fair representation and engagement of relevant stakeholders, from all levels of government, industry, emergency management, and security. Similarly, for successful information sharing to occur, there must be established mechanisms or channels to reach stakeholders regularly during blue-sky conditions and the various stages of an incident. Sharing information can take many forms, including training events, briefings, e-mail alerts, conference calls, meetings in secure locations to discuss classified materials about specific threats or hazards, and documents and forums that encourage sharing lessons learned.

CIPAC is the primary framework to facilitate voluntary collaboration and information sharing within and across critical infrastructure sectors through the public-private partnerships and councils outlined above. Beyond CIPAC, however, there are several other programs or venues that facilitate collaboration and information sharing. This section will introduce two of these programs. First, information sharing and analysis centers enable information sharing between the government and private sector by use of a sector-based model, meaning that organizations within a certain critical infrastructure sector (or a specific segment within a sector) join to share information.[34] For organizations that do not fit neatly within an established sector or that have unique information requirements, a more viable option may be to join one of a number of information sharing and analysis organizations. Focused on gathering, analyzing, and disseminating cyber threat information among members, these organizations offer a more flexible approach to self-organized information sharing among specific communities of interest,

32. DHS, *NIPP 2013*, 7.

33. DHS, *NIPP 2013*, 7.

34. "Information Sharing and Awareness," CISA (website), n.d., accessed January 26, 2022, https://www.cisa.gov/information-sharing-and-awareness.

such as legal or accounting firms that support clients across several critical infrastructure sectors.[35]

Given its collective experience since the mid-1990s, the critical infrastructure community has learned important lessons regarding information sharing, that over time have been incorporated into US CISR policy and practice. Some of these key lessons are: (1) to include relevant CISR stakeholders while protecting owner and operator information, (2) to share actionable threat information so owners and operators can take appropriate action, (3) to encourage and practice reciprocal and multidirectional information sharing, and (4) to pursue methods to disseminate threat information safely and more broadly, especially with private-sector companies without appropriate security clearances.[36] Since this list of collaboration and information-sharing best practices in the United States is not exhaustive, it is helpful to examine the more thorough analysis and discussion of information and intelligence-sharing principles, frameworks, and best practices outlined in chapter 11.

Moving Forward: Sustaining CISR Success for the Long Term

As the United States has learned over nearly 30 years of experience, establishing and implementing a demonstrably effective national CISR policy is one of the most difficult things a nation can do. Even with an established, strong CISR policy, culture, and practices, the United States depends on the continuation of strong public-private partnership, the transfer of institutional knowledge, and the investment in its human capital to sustain and mature this posture in the future.

Training, education, and exercises are fundamental to the long-term success of the national CISR posture and must include government officials, infrastructure owners and operators, first responders, and the general public where appropriate. Training can focus on general concepts, best practices, or specific topics and should be available in many different forms to ensure the broadest reach, including means such as instructor-led courses, web-based independent-study courses, and written guidance and job aids. The DHS, for example, currently offers a range of training topics for use by government agencies or private-sector entities, including sector-specific best practices, dealing with insider threats or active-shooter scenarios, supply-chain risk

35. "Information Sharing and Analysis Organizations (ISAOs)," CISA (website), n.d., accessed January 26, 2022, https://www.cisa.gov/information-sharing-and-analysis-organizations-isaos.

36. CISA, *Critical Infrastructure Security*, 12.

management and third-party dependency, industrial control systems and operational technology, and incident management and response.[37]

Similarly, exercises can reinforce training and education by providing scenarios to apply knowledge and skills in an operational setting, simulate real-world threats and appropriate response options, and strengthen trust and understanding in relationships within and among participating organizations. Using a variety of models—including workshops or seminars, tabletop exercises, rehearsals of key plans, functional exercises, and full-scale exercises—will help the United States (or any country) maintain and grow a culture of continuous improvement in its CISR posture so the nation is prepared for the current and next generation of threats to its critical infrastructure.

To sustain success in established CISR programs, organizations should promote the program and periodically assess and evaluate it. When it comes to promoting the CISR program, regardless of the level at which it exists, the US approach highlights the importance of outreach and awareness. While CISR programs engage their various stakeholders—including private-sector companies, local governments, and citizens—in different ways, ensuring they understand the risks, have sufficient information, and can make decisions regarding risk mitigation and management with confidence are crucial elements to ensure success.[38] Using social media, web-based training, public media outlets, and conference presentations are just a few ways to reach a broad range of stakeholders. The DHS-led campaign, If You See Something, Say Something®, is a useful example. This effort has successfully reached beyond personnel directly involved with critical infrastructure operations to include entire communities and increased their situational awareness. See chapter 11 for a detailed discussion of this initiative.

Similarly, periodic evaluation of existing CISR programs is essential to ensure they adapt to emerging threats and apply appropriate measures to enhance security and resilience. Since CISR programs can involve representatives from multiple sectors, several levels of government, and owners and operators from different facilities or systems, they can also be difficult to evaluate effectively. The US CISR framework recognizes and seeks to balance two competing imperatives for program assessment.[39] First, CISR programs should have measurements that are simple

37. CISA, *Critical Infrastructure Security*, 19.

38. CISA, *Critical Infrastructure Security*, 21.

39. CISA, *Critical Infrastructure Security*, 20.

to conduct consistently over time, document and compare actual performance across sectors and regions, and identify shortcomings or gaps in performance along with corrective measures to address them. Second, they should have customized performance metrics that can work in any reporting situation. Thus, a blend of performance measures common to each sector, along with nuanced metrics for the subsector level, can help meet these needs.

Together, these steps of education and training the human work force, promoting the program to a broader population, and developing suitable frameworks for assessing and evaluating performance contribute to CISR programs that can succeed in the future as new threats, risks, and vulnerabilities emerge. While the US CISR framework endorses these steps, they are not unique to the American situation. In many respects, the EU CISR framework faces similar challenges to those in the United States and undertakes similar measures to address them. Whereas the US CISR policy relies on coordination across various levels of government and jurisdictions and strong partnerships between the public and private sectors, the EU must coordinate CISR efforts across its 27 member states, each with its own blend of critical infrastructure facilities, public-private partnerships, and questions of national sovereignty. The next section will explore the CISR policy and practices used in the EU, which not only affects the member states but also other European nations that are not EU members but are located adjacent to or within the EU's borders.

EU CISR Policy Framework

A series of terrorist attacks in the early years of the twenty-first century served as the primary motivation to develop a policy for enhanced collective CISR posture in Europe. In 2004, the EU launched its first joint CISR measures in the wake of the bombings of the train network in Madrid, with the 9/11 terrorist attacks still fresh in the West's memory and just one year before the attacks against the London underground system. See chapter 7 for its detailed case studies of the Madrid and London railway attacks. These attacks convinced the EU, both at the institutional and member state levels, of the importance of establishing common rules, mechanisms, and tools to foster better governance, management, and protection of national and European critical infrastructures (NCI and ECI, respectively).

The EU's objective is not to interfere in matters of national security, which remain the exclusive responsibility of each member state. Instead, EU CISR policy fosters the security and resilience of critical

infrastructures in the EU through a program that harmonizes efforts across the member states and enables a more mature CISR posture overall. In 2004, the maturity of member states' individual CISR policies and practices was very uneven. A small number of countries, such as France, Germany, and the United Kingdom, already had a national framework in place at this time, while the vast majority of member states relied on a basic set of rules or very embryonic approaches. Even though most countries had various elements of CISR programs within their national legislation, they were very fragmented and lacked long-term objectives consistent with the unique nature of European critical infrastructure. For instance, national CISR programs overlooked measures to prevent trans-boundary externalities and domino effects, which are some of the drivers of the EU's current approach.

Given these conditions, the EU established the European Programme for Critical Infrastructure Protection (EPCIP) in 2006 and, since then, this program has served as the foundation for planning, executing, and consolidating most European CISR activities. In response, EU institutions—particularly the European Commission—have played key roles in setting the conditions for the overall improvement of CISR capabilities across Europe by fostering a risk-based approach. Using risk assessment and management as the foundation for CISR efforts has now become so ingrained that this mentality influences even the most recent directives and regulations addressing cybersecurity, such as the 2019 EU regulation that updated the mandate of the EU Agency for Network and Information Security (ENISA) to address emerging risks and threats in the cyber domain.[40]

The EPCIP has improved the security and resilience of critical infrastructures in both the public and private sectors by persuading member states to establish or renew their national frameworks and improve their cooperation with private-sector operators and critical infrastructure stakeholders in neighboring states. Although the EPCIP has been a successful program, the objectives of the plan have taken longer to materialize than originally envisaged. One example that illustrates this ambitious timeline is the EU's commitment to review Directive 114/08/EC—which called on member states to complete an assessment to identify and designate facilities or sectors

40. *Regulation (EU) 2019/881: On ENISA (the European Union Agency for Cybersecurity) and on Information and Communications Technology Cybersecurity Certification and Repealing Regulation (EU) No 526/2013 (Cybersecurity Act)*, European Parliament and the Council, April 17, 2019, https://eur-lex.europa.eu/eli/reg/2019/881/oj.

that qualify as critical infrastructure—four years after its entry into force.[41] The review planned for 2012 never took place because some member states had only recently transposed the directive into their national legislation and still needed to complete the operative phase of identification, designation, and protection of critical infrastructure. Directive 114/08/EC is still in force, though it will likely be repealed by 2023 and replaced by the directive on critical entities resilience introduced in late 2020.[42]

The harmonization of numerous national approaches toward the identification, designation, and protection of both NCIs and ECIs has proven to be quite challenging. This difficulty is due not only to the various levels of CISR posture and capability among member states, but also to these 27 individual countries having processed the security objectives outlined by the EU through their respective cultures, governance models, histories, economies, and preexisting priorities. At an institutional level, the EU is governed by the principle of subsidiarity, which ensures member states retain the ability to make decisions and act in areas in which the EU does not have exclusive competence.[43] These circumstances have deeply influenced the way each country has responded to the common security agenda outlined in the EPCIP. In this way, the pursuit of a strong European CISR posture presents unique challenges when compared to the United States given the substantial difference in their respective governance models.

To date, the EU institutions, member states' national governments, critical infrastructure owners and operators, academia, and research centers have worked together to increase baseline levels of security and resilience across Europe. The slow and steady progress the EU and its member states have achieved in improving awareness and developing more mature CISR policies and practices reflects a mentality of continuous improvement, which has led the EU to adopt several additional measures and sector-specific policies beyond the EPCIP. When considered in a holistic manner, these advances offer an encouraging snapshot of progress from a CISR program that was initially focused on protection and

41. *Council Directive 2008/114/EC: Identification and Designation of European Critical Infrastructures and the Assessment of the Need to Improve Their Protection,* Council of the European Union, December 8, 2008, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:345:0075:0082:EN:PDF.

42. Sebastijan R. Maček, "EU Members Agree on Resilience of Critical Infrastructure," Euractiv (website), December 21, 2021, https://www.euractiv.com/section/eu-council-presidency/news/eu-members-agree-on-resilience-of-critical-infrastructure/.

43. "The Principle of Subsidiarity," European Parliament (website), October 2021, https://www.europarl.europa.eu/factsheets/en/sheet/7/the-principle-of-subsidiarity.

characterized by isolated national approaches to a framework that prioritizes resilience and is increasingly harmonized, coordinated, and comprehensive across EU member states. Figure 10-3 below illustrates the progression of EU CISR policy since 2004.

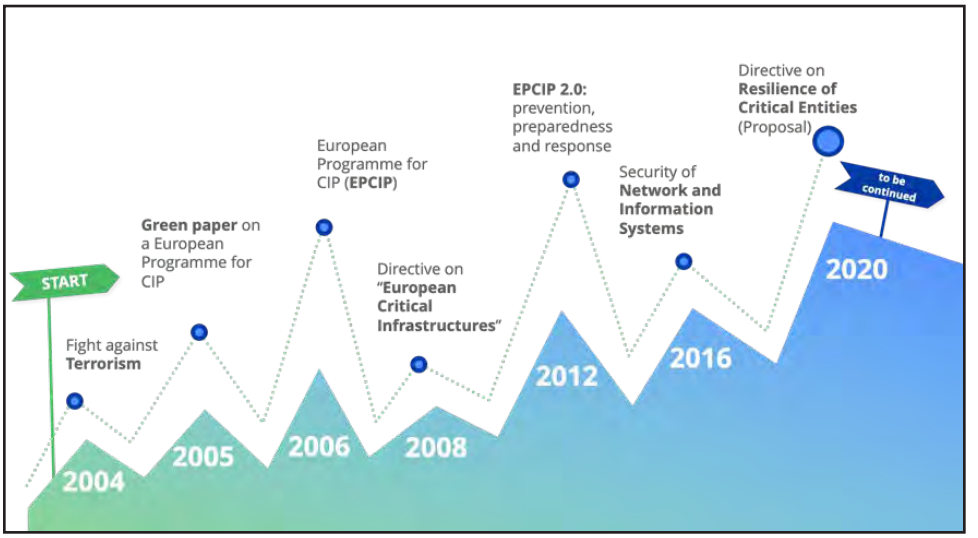


Figure 10-3. Milestones in EU CISR policy development (2004–20)

The following sections will provide a macro-level description of the phases and milestones depicted in figure 10-3, which have characterized the evolution of the EU's CISR policy framework over the last two decades. The section will focus on the main milestones and will not discuss sector-specific measures, such as the policy, directives, and regulations in key areas like port security or civil aviation.[44]

2004: Embryonic Stage Motivated by Fight against Terrorism

The terrorist attacks of Madrid, which took place in 2004, were the spark that ignited the launch of a joint EU program on critical infrastructure protection (CIP). These tragic events unveiled the fragility of daily life in European society, including the strategic assets and critical infrastructures that enable its prosperity and provide essential services each day. They also provided the context and motivation for the EU at an institutional level to pursue several initiatives that would eventually form the foundation

44. See *Directive 2005/65/EC: Enhancing Port Security*, European Parliament and the Council (website), October 26, 2005, https://eur-lex.europa.eu/eli/dir/2005/65/2019-07-26; and *Regulation (EC) No 300/2008: Common Rules in the Field of Civil Aviation Security and Repealing Regulation (EC) No 2320/2002*, March 11, 2008, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32008R0300.

and objectives of the EPCIP. Several important communications by the European Council and the EU Commission throughout 2004 illustrate this new trajectory.

From the perspective of the Council, two vital documents capture the essence of the early and developing nature of CIP policy. First, the presidency conclusion from the Brussels European Council sessions on July 17–18 directed the Council and the Commission to "pursue with determination the objective of building a shared area of freedom, security and justice" for the coming years and to develop proposals for a new program that would achieve this strategic goal.[45] At the same time, the Council also charged the Commission to work with the Council in preparing an overall strategy to enhance the protection of critical infrastructures and integrate the fight against terrorism fully into EU external relations policy. The second important event came four months later, when the Council published an update to the EU Plan of Action on Combating Terrorism. The highlights of this plan are listed below.[46]

- Enhance international efforts and will to combat terrorism.
- Reduce the access of terrorists to financial and other economic resources.
- Maximize the capacity within EU bodies and member states to detect, investigate, and prosecute terrorists, prevent terrorist attacks, and deal with the consequences of a terrorist attack.
- Secure international transport and ensure effective border control systems.
- Address factors that contribute to support for and recruitment into terrorism.
- Target actions under EU external relations toward priority third countries that need to enhance their commitment and capacity to combat terrorism.

Similarly, the European Commission published four specific communications to the Council and the European Parliament

45. Council of the European Union, *Presidency Conclusions—Brussels, 17 and 18 June 2004* (Brussels: European Council, 2004), 2, https://data.consilium.europa.eu/doc/document/ST-10679-2004-REV-2/en/pdf.

46. Council of the European Union, *EU Plan of Action on Combating Terrorism—Update* (Brussels: Council of the European Union, 2004), 2, https://data.consilium.europa.eu/doc/document/ST-14330-2004-REV-1/en/pdf.

on October 20, 2004. Together, these communications demonstrated the Commission's commitment to moving beyond individual, national approaches to a more structured, wide-ranging, and collective EU framework that could facilitate a timely, adequate, and more coordinated response to all terrorist scenarios. Together, these communications introduced several priority topic areas, such as the availability of common alerting systems for prompt and effective communications with citizens and measures to improve the exchange of information, strengthen transparency, and enhance the ability to trace terrorists' financial transactions.[47]

Perhaps the most important of these communications, *Critical Infrastructure Protection in the Fight against Terrorism*, clarified the definition of critical infrastructure, designated specific sectors as critical infrastructure, and outlined the fundamental concepts and requirements that became the foundation for the EPCIP. This communication defined critical infrastructures as "those physical and information technology facilities, networks, services and assets which, if disrupted or destroyed, would have a serious impact on the health, safety, security or economic well-being of citizens or the effective functioning of governments" in the various EU member states.[48] The communication also identified nine sectors as critical infrastructure: (1) energy installations and networks, (2) communications and information technology, (3) finance, (4) health care, (5) food, (6) water, (7) transport, (8) the production, storage, and transport of dangerous goods, and (9) essential government services and functions.[49]

Although this initial list of critical infrastructures is quite comprehensive, the Commission determined that further refinement and definition were necessary. The Commission specifically directed member states to identify and designate which critical infrastructures qualify as national (NCI) and charged EU institutions to determine which are European (ECI) by the end of 2005. Identifying NCIs and ECIs was a critical first step because it allowed the EU—in line with the principle of subsidiarity and recognizing the highly connected and interdependent nature of these services and networks—to focus its efforts on those sectors that most impacted the functionality of critical infrastructure in the member states. To guide

47. Commission of the European Communities, *Communication from the Commission to the Council and the European Parliament on the Prevention of and the Fight against Terrorist Financing* (Brussels: European Commission, 2004), 3, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52004DC0700&from=EN.

48. Commission of the European Communities, *Communication from the Commission to the Council and the European Parliament: Critical Infrastructure Protection in the Fight against Terrorism* (Brussels: European Commission, 2004), 3, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52004DC0702&from=GA.

49. Commission, *Critical Infrastructure Protection*, 4.

EU institutions and member states in designating NCIs and ECIs, the Commission recommended the three factors outlined below for consideration.[50]

- Scope. The geographic area—international, national, or local, for example—affected by the loss or unavailability of a critical infrastructure.
- Magnitude. The degree of the impact or loss of a critical infrastructure can be assessed as none, minimal, moderate, or major. Useful criteria to evaluate potential magnitude include the impacts on: the general population, the economy, the environment, other critical infrastructure sectors or facilities, and the government's ability to function properly.
- Effects of time. At what point the loss of a critical infrastructure could have serious impact: immediately, in a matter of days or weeks, or other.

Finally, this crucial communication justified the need for an EU-wide CIP program consisting of a common framework, but with clear responsibilities for the member states and the EU, respectively. Citing the impossible task of trying to protect all possible infrastructures the member states might identify as critical, the Commission asserted that, at an institutional level, the EU must prioritize the protection of infrastructures that have a transboundary effect, leaving all others under the purview of the member states.[51] This focus on transnational networks and cross-border connectedness was at the core of EPCIP from the start, and it remains one of the enduring elements and guiding principles of the EU CISR framework. On this point, the Commission also outlined three metrics by which it would measure the success of an eventual EPCIP: (1) member states identifying and establishing inventories of critical infrastructures according to EPCIP priorities, (2) businesses collaborating within sectors and with government to share information and reduce the likelihood of major incidents disrupting critical infrastructures, and (3) an EU-wide common approach to tackling the security of critical infrastructures through public-private cooperation.[52] These three criteria constitute the initial pillars upon which EU institutions

50. Commission, *Critical Infrastructure Protection*, 5.

51. Commission, *Critical Infrastructure Protection*, 7.

52. Commission, *Critical Infrastructure Protection*, 9.

and member states established the EPCIP and were the basis for their intensive discussions on how to mature the program.

2005: From the Fight against Terrorism to an All-hazards Approach

In the timeline of events, 2005 is a key year because it is when the EU adopted an all-hazards approach as the basis for an EPCIP and its efforts to protect critical infrastructures from the wide variety of threats that could disrupt operations or destroy facilities and assets. The matter of a strong, collective CISR posture has become more prominent in the European policy agenda as the fear emerged that negative effects caused by the disruption, failure, or destruction of a critical infrastructure in a single member state could spread to others as well. While the key initiatives of 2004 introduced the mandate to prepare a strategy to protect critical infrastructures and eventually establish an EPCIP, the Commission's immediate actions aimed at gathering a critical mass of stakeholders involved in the lifecycle of critical infrastructures. The Commission organized two seminars in 2005 during which member states shared the status and progress of their CIP programs and exchanged information with private-sector representatives on how to better define the respective competencies and domains of interest.

As a result of the discussions at these two seminars, the Commission published the *Green Paper on a European Programme for Critical Infrastructure Protection*, which included the analysis of numerous policy options for how to respond to the Council's 2004 request to establish an EPCIP. As set out in the initial stages of 2004, focus of an EPCIP would be to achieve an adequate level of protection for all the critical infrastructures with priority focus on those which, if disrupted, would cause severe impacts across multiple EU member states. Perhaps most importantly, the paper argued that the optimal way to strengthen critical infrastructure in the EU was to establish a common EPCIP framework and facilitate the exchange of best practices and ways to monitor compliance.[53]

In the green paper, the Commission also included two essential elements of a future EPCIP: a potential definition for ECI and a continued emphasis on the importance of assessing transboundary effects. First, the paper retains the definition of critical infrastructure as those resources, services, facilities, networks, and assets, which, if disrupted or destroyed would have serious impacts on health, safety, security, economic, or social well-being.

53. Commission of the European Communities, *Green Paper on a European Programme for Critical Infrastructure Protection* (Brussels: EU Commission, 2005), 2–5, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52005DC0576&from=EN.

When it comes to qualifying areas that should be European critical infrastructure, the Commission has presented two options to consider when designating ECI: those which if disrupted or destroyed (1) would have a serious impact on two or more member states or (2) would involve three or more member states.[54] The emphasis on transboundary effects constituted the main driver of the future EU policy as a part of the definition and a requirement for identifying and designating ECIs.

Given the transboundary and interdependent nature of critical infrastructures across the EU, the Commission proposed that a common EPCIP framework—though primarily focused on ECIs—should also form the basis for how member states identify, designate, and protect their NCIs. The Commission recommended this top-down approach from the standpoint of providing critical infrastructure owners and operators a more simplified and efficient framework and serving the best interests of the individual member states and the EU as a whole.[55] Although the Commission suggested member states had the flexibility to create national CIP organizations that could apply additional, more demanding measures than those outlined in an EPCIP, the core issue was the application of the principle of subsidiarity. While member states generally endorsed a collective EU approach, they also had a strong desire to retain control over those critical infrastructures in their respective national borders that would be designated as ECIs.

While the green paper's policy recommendations initiated this tension over EU-level management of certain critical infrastructures and national control of others, the EU *Council Directive 2008* would later provide more clarity. On this point, the council directive affirmed the member states' sovereignty to designate which critical infrastructures located within their national borders could also be designated as an ECI and clarified the Commission's role to support the member states in this process.[56] Similarly, many other elements and best practices introduced in the green paper would later form the foundation of the EPCIP in 2006 and inform the guidance outlined in the *Council Directive 2008*, which today constitutes the core element of the EPCIP.

54. Commission, *Green Paper*, 6–7.

55. Commission, *Green Paper*, 9–10.

56. *Council Directive 2008/114/EC: Identification and Designation of European Critical Infrastructures and the Assessment of the Need to Improve Their Protection*, European Council, December 8, 2008, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0114&from=EN.

2006: EU Formally Creates EPCIP

Following the consultations and discussions triggered by the green paper, the *Communication from the Commission on a European Programme for Critical Infrastructure Protection*, published in 2006, formally created the EPCIP as the framework to protect critical infrastructure across the EU, outlined the principles, procedures, and tools to implement it, and promised a council directive to ensure the achievement of the joint protection objectives.[57] Although it recognized terrorism as the priority threat to critical infrastructure, the basis of the EPCIP was an all-hazards approach. As the framework for EU-wide critical infrastructure protection, the EPCIP consisted of several essential elements to guide future efforts.[58] The list below contains the key concepts put forth in the EPCIP:

- A procedure to identify and designate ECIs.
- A common approach to assess and enhance protection of ECIs.
- Support for member states to identify, designate, and protect their NCIs.
- Contingency planning to minimize the effects of disruptions of ECIs or NCIs.
- An external dimension to assess impacts of CIP outside of EU borders.
- Financial measures to provide for initiatives and efforts related to CIP.
- Several measures to further develop and implement the EPCIP.

Among the measures to implement the EPCIP were several ways to exchange best practices, share information, and improve dialogue between relevant CIP stakeholders. These measures consisted of expert groups at the EU level, CIP information-sharing processes to promote trust and protect sensitive information, a process to identify and analyze interdependencies based on geography and sector, and the critical infrastructure warning information network (CIWIN). The framework also included the EPCIP Action Plan

57. Commission of the European Communities, *Communication from the Commission on a European Programme for Critical Infrastructure Protection* (Brussels: EU Commission, 2006), 2–3, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52006DC0786&from=EN.

58. Commission, *European Programme for Critical Infrastructure Protection*, 3–4.

based on three main work streams: (1) to provide a strategic platform for overall EPCIP coordination and cooperation across the EU institutions and member states via the EU CIP contact group, (2) to enhance the protection of ECIs by reducing their vulnerability, and (3) to support member states in all efforts regarding NCIs.[59] One of the strongest elements of the action plan was the creation of the EU CIP contact group. This group, comprised of one CIP representative from each member state charged to coordinate national CIP issues, provided a forum chaired by the EU Commission in which member states could discuss issues and make decisions with each other and with the Council and Commission.[60]

Finally, the EPCIP made progress on the contentious issue of NCIs by clarifying that the responsibility for NCIs rests with member states and NCI owners and operators, and that the Commission would support member states as requested. The EPCIP encouraged each member state to draw up a national CIP program, based on the process and approach used for ECIs, to protect the NCIs located in its respective national territory. According to the EPCIP, the national programs should address: (1) the identification and designation of NCIs based on the geographic extent of damages and the severity of consequences resulting from the disruption or destruction of these infrastructures, (2) the identification of geographic and sectoral interdependencies, (3) the establishment of dialogue with the owners and operators responsible for protection of NCIs, and (4) the development of contingency plans as needed.[61] To some extent, elements of these national CIP programs are still under development since, under its practice of continuous improvement, the EU aims to consolidate collective results and then set new objectives and thresholds for success as the overall maturity of CIP allows.

2008: Identifying, Designating, and Protecting ECI

The year 2008 marks the next critical milestone in EU CISR policy. In this year, the EU Council Directive established the procedure to identify and designate ECIs and outlined a common approach to assess how to improve the protection of those ECIs. For clarity, the Council Directive defined an ECI as an infrastructure located in a given member state, the disruption or destruction of which would have a significant impact on at least two member states in terms of several crosscutting criteria, such as those that result

59. Commission, *European Programme for Critical Infrastructure Protection*, 5, 10–13.

60. Commission, *European Programme for Critical Infrastructure Protection*, 4.

61. Commission, *European Programme for Critical Infrastructure Protection*, 7.

from cross-sector dependencies on other infrastructure sectors or systems.[62] Figure 10-4 illustrates the step-by-step procedure for the identification and designation of ECIs and the common approach for their protection.

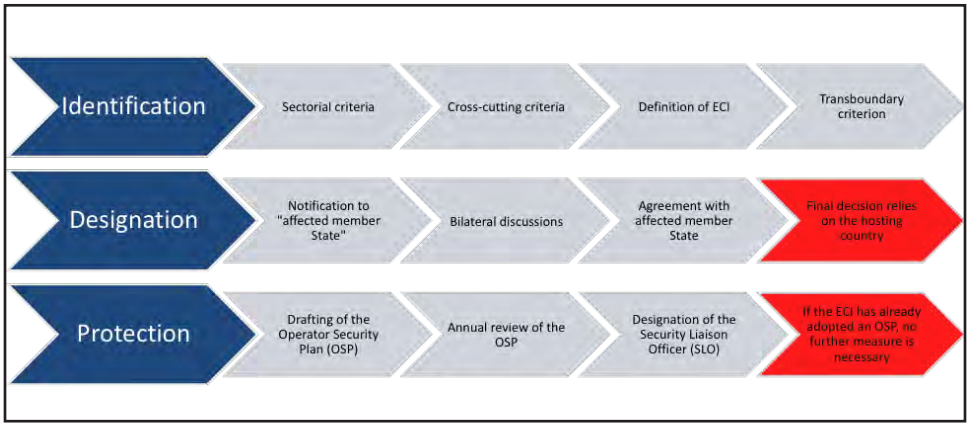


Figure 10-4. Process for identifying, designating, and protecting ECIs

As anticipated, the Council determined to use the energy and transport sectors for the purpose of implementing the directive because these two sectors involved the most substantial connections, interdependencies, and transboundary effects—the key concepts established throughout the evolution of EU CISR policy since 2004. Concerning the sectoral criteria for identifying ECIs, table 10-1 describes the directive's focus on the energy and transport sectors and their respective subsectors.[63]

62. *Council Directive 2008/114/EC.*

63. *Council Directive 2008/114/EC.*

Table 10-1. ECI sectors and subsectors

Sector	Subsector	
Energy	Electricity	Infrastructures and facilities to generate and transmit electricity (supply of electricity)
	Oil	Oil production, refining, treatment, storage, and transmission by pipelines
	Gas	Gas production, refining, treatment, storage, and transmission by pipelines, including LNG terminals
Transport	Road	
	Rail	
	Air	
	Inland waterways	
	Ocean and short-sea shipping and ports	

When applying the sectoral criteria to identify potential ECI, the directive foresees that member states would need to consider alternatives—those infrastructures or services that could function as backup options (such as alternate ports or LNG terminals)—if proposed ECIs are not available. The lack of a well-defined criterion for appropriate alternatives has effectively limited the number of identified ECIs and served as a work-around to avoid a designation.

After the sectoral criteria outlined above, the next step in identifying ECIs is to apply several crosscutting criteria. These criteria include: (1) casualties by the potential number of fatalities or injuries, (2) economic effects assessed by the significance of potential economic loss and/or degradation of products or services, including environmental effects, and (3) effects on the public in terms of the impact on public confidence, physical suffering, and disruption of daily life, including the loss of essential services.[64] Finally, the last steps of the identification process require member states to apply the definitions of critical infrastructure and ECI and the transboundary criterion to determine which infrastructures, if disrupted or destroyed would have a significant impact on at least two member states.

The formal ECI designation process is more political in nature than technical since it implies the cooperation of the EU member states with the member state on whose territory the ECI is located making the final decision. Such procedure in the initial phase of the application of the directive led

64. *Council Directive 2008/114/EC.*

to several failures to designate ECIs. These failures were due to several factors, including: (1) insufficient cross-border CIP cooperation between member states, (2) certain member states' unwillingness to put additional pressure on critical infrastructure owners already engaged in national security projects, and (3) the requirement for designated ECI to prepare an operator security plan (OSP) and appoint a security liaison officer.

Among these reasons, the need for proper security management through the OSP and an appointed security liaison officer are essential elements of the process to protect ECI. Based on a risk-driven approach derived from basic international standards on risk management, the OSP consists of three key steps: (1) identify the critical infrastructure's vital processes and assets that ensure functionality, (2) conduct a risk analysis based on threats, vulnerabilities, and possible impacts, and (3) develop, prioritize, and employ control measures to mitigate risk and enhance overall protection. Figure 10-5 provides an overview of the purpose and key steps of the OSP.

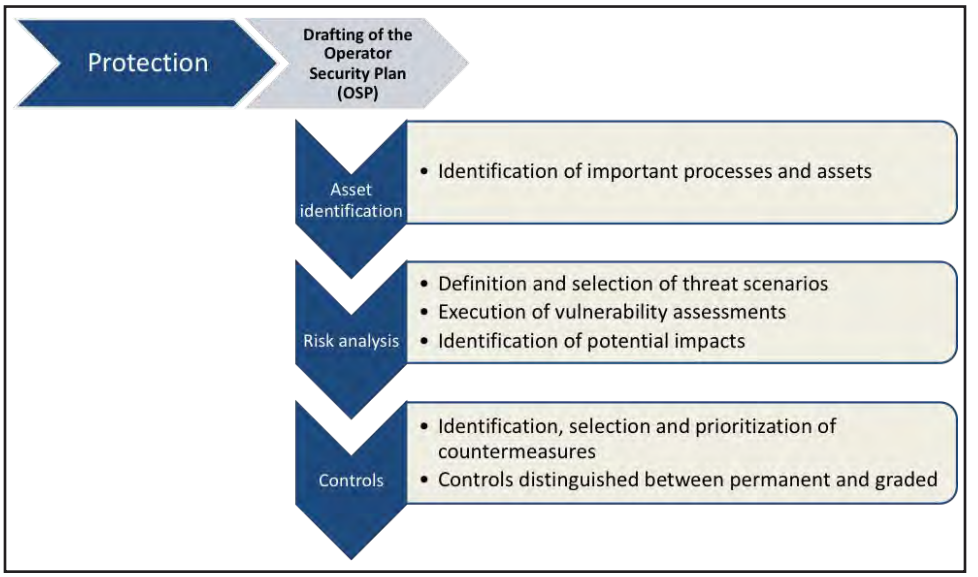


Figure 10-5. Approach for developing an operator security plan

2013: EPCIP 2.0—A New Approach

Since the inception of its first collective CIP efforts, the EU's approach has been one of steady progress aimed at incrementally improving the framework over time and highlighting the respective focus areas for the EU, member states, and critical infrastructure owners and operators. Although the initial steps of the European journey embraced terrorism and

protection as the main priorities, relevant stakeholders eventually came to recognize the importance of topics like the all-hazards approach, resilience, and cybersecurity. When the EU began pursuing collective CIP efforts in 2004, the member states did not share a common framework or approach, and their CIP capabilities were at very different levels. By taking these realities into consideration, the EU calibrated its efforts to avoid overwhelming member states that were updating their respective CIP frameworks, strategies, policies, and procedures—sometimes from scratch.

By 2013, EU institutions and member states recognized they had collectively achieved an acceptable level of CIP maturity and should consider the introduction of new EU-wide policies and objectives to enhance CIP. In this spirit of continuous improvement, a series of EU-level discussions led to the EU advocating the initial elements of a new concept: the resilience of critical infrastructure. The embrace of resilience as a formal concept first appears in the *EU Commission's Staff Working Document (SWD) 318*, which focused on a new approach and more practical implementation of the EPCIP. *SWD 318* began with the premise that ensuring a high degree of protection of EU infrastructures and making them more resilient to all hazards and threats can minimize the consequences of loss of services across Europe.[65]

In addition to the topic areas previously covered under the EPCIP, *SWD 318* called on member states to consider and integrate prevention, preparedness, and response into their national approaches. These three areas formed the pillars upon which the Commission desired to build future initiatives after consolidating the achievements to date under the EPCIP. These initial elements of the path toward a resilience-based approach provided a new way for the member states to interpret and implement the EU's framework for critical infrastructures. In this way, *SWD 318* effectively marked the EU's shift from critical infrastructure protection (CIP) to security and resilience (CISR) because it asked member states to extend their efforts beyond only preventing attacks to also taking the necessary steps to prepare for such attacks and develop the capability to respond to and recover from them as quickly and smoothly as possible.

65. European Commission, *Commission Staff Working Document on a New Approach to the European Programme for Critical Infrastructure Protection: Making European Critical Infrastructures More Secure* (Brussels: European Commission, 2013), 2, https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2013)318&lang=en.

2016: Directive on Network and Information Security

Consistent with its continuous improvement mentality, the EU took a major step forward to improving the cybersecurity of critical infrastructures in 2016 when the European Parliament and the Council published its directive on network and information security (NIS). By 2016, member states had already been focusing on cybersecurity issues and adopting measures to mitigate risks in the cyber domain in their national frameworks. After all, the 2008 Council Directive anticipated the need to include sectors other than just energy and transport, specifically citing the information and communication technology–sector as one of these areas of emerging interest.[66] Similar to the 2008 Council Directive on ECIs, the central objective of the 2016 NIS directive was to harmonize the identification and designation phases, including the path to achieve better security of network and information systems and the provision of mechanisms to foster qualified and effective cooperation among member states.[67] The NIS directive, however, applied a dramatic change to understanding and enhancing CISR because it proposed an approach focused on essential services instead of critical infrastructures. The rationale for this decision was to differentiate the pillars of physical security (critical infrastructures) and cybersecurity (essential services) while also signaling clearly the need to focus on protecting the services that rely on network and information systems.

As under the ECI Directive, the designation phase for network and information systems relies in part on a sectoral approach. While the ECI Directive only covered the energy and transport sectors, the NIS Directive expanded its focus to include seven sectors: energy, transport, water, banking, financial-market infrastructures, health care and digital infrastructures. The wider range of sectors included in the NIS Directive indicates how member states' commitment to CISR efforts has grown since the Council published the ECI Directive in 2008.

The NIS Directive aims to improve the overall level of security of network and information systems across the EU. In practice, this objective means network and information systems can resist any action that "compromises the availability, authenticity, integrity, or confidentiality" of stored, transmitted, or processed data as well as any services those systems may offer or enable.[68]

66. *Council Directive 2008/114/EC.*

67. *Directive (EU) 2016/1148: Measures for a High Common Level of Security of Network and Information Systems across the Union*, European Parliament and the Council, July 6, 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.194.01.0001.01.ENG&toc=OJ:L:2016:194:TOC.

68. *Directive (EU) 2016/1148*, Article 1.

To achieve this objective, the NIS Directive addressed these vital areas: (1) member states' preparedness through means such as developing a national NIS strategy or creating computer security incident response teams (CSIRT), (2) enhanced cooperation among the member states via the Cooperation Group, (3) an improved EU culture of NIS security, (4) the adoption of security measures to mitigate risks to systems, and (5) the development of a mechanism to notify national authorities of NIS security incidents.[69]

Given the NIS Directive's focus on essential services, it proposes a procedure to identify which entities meet these qualifications. Entities that qualify as operators of essential services are those that: (1) operate in one of the seven sectors outlined above, (2) provide a service essential for the maintenance of critical societal and/or economic activities, (3) depend on network and information systems to provide these services, and (4) would face significant disruptive effects on their ability to provide these services in the event of a NIS security incident.[70] To define the term *significant disruptive effects*, the directive outlines several cross-sectoral factors to consider, such as the number of users who rely on the service, the extent to which other sectors are dependent upon the service, the impact that incidents could have on the economy, society, or public safety, and the geographic area that could be affected by an incident. Additional considerations are the entity's market share of the service provided and its importance for maintaining a sufficient level of the service when taking into account alternate means.[71]

Beyond outlining the process for identifying and designating operators of essential services, the NIS Directive also establishes several security requirements. First, operators of essential services should take appropriate and proportionate measures to manage the risks to their NIS security. Next, operators should take steps to prevent and minimize the impact of incidents affecting their NIS security and seek to ensure continuity of these essential services. Lastly, they should notify the competent national authority or the CSIRT of incidents that seriously impact the continuity of their services and provide information that enables these authorities to determine any cross-border impacts that may have resulted.[72] During the preparation of the final text of the NIS Directive, stakeholders intensely debated the requirement for incident notification

69. *Directive (EU) 2016/1148*, Article 1.

70. *Directive (EU) 2016/1148*, Article 5(2).

71. *Directive (EU) 2016/1148*, Article 6.

72. *Directive (EU) 2016/1148*, Article 14.

because it effectively created a new kind of social contract between operators and member-state governments. Such a notification mechanism cannot achieve its full potential without the cooperation of operators and the appropriate response measures governments undertake to minimize the effects of the incident. Governments can improve prevention, preparedness, and overall security only when they have timely access to vital information regarding incidents.

In this way, the NIS Directive proposes an unprecedented effort, especially in its establishment of the Cooperation Group and the CSIRT network. The Cooperation Group—comprised of representatives from the member states, the Commission, ENISA, and other relevant stakeholders as required—serves to exchange NIS security information and best practices, discuss levels of capability and preparedness, and provide strategic guidance to the CSIRT network.[73] Similarly, the CSIRT network—consisting of member states' CSIRTs and the EU's computer emergency response team, with the Commission in an observer role and ENISA as the secretariat—exists to deepen confidence and trust between member states and promote effective operational cooperation in NIS security.[74] Together, these two entities take pivotal actions to address both the strategic and tactical domains of cybersecurity while also enabling prompt and effective information sharing across the EU and between the public and private sectors.

The promulgation of the NIS Directive represents the EU's first full cycle of efforts to address the physical security of critical infrastructures and the cybersecurity of the essential services they provide. Since 2016, member states' commitment in these domains has been growing consistently. The EU's pursuit of an impact assessment and public consultation on the NIS Directive, despite its relatively short existence, is one example that demonstrates this commitment.[75] The next section will build on this theme of continuous improvement in EU CISR policies and practices.

2020: Proposal for Directive on Resilience of Critical Entities

December 16, 2020, should be included among the landmarks of the EU's CISR history because the Commission published two proposals for new CISR-related directives on this date. With the release of these

73. *Directive (EU) 2016/1148*, Article 11(3).

74. *Directive (EU) 2016/1148*, Article 12.

75. "Cybersecurity—Review of EU Rules on the Security of Network and Information Systems," European Commission (website), n.d., accessed on October 12, 2021, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12475-Cybersecurity-review-of-EU-rules-on-the-security-of-network-and-information-systems_en.

two proposals, the EU opened a new cycle aimed at drastically improving security in the physical and cyber domains of critical infrastructures and the essential services they provide. First, the Commission submitted a proposal to the European Parliament and the Council to repeal the original NIS Directive and adopt an updated version—the so-called NIS 2.0—that brings further improvements, especially in ways to facilitate cooperation.[76] Second, the proposal for a directive on the resilience of critical entities proposed a recalibration of the focus on physical security, which was still governed by the 2008 ECI Directive and therefore considered obsolete and too limited in scope since it only covered the energy and transport sectors.[77] With this double proposal, the Commission sought to harmonize its efforts in the domains of physical security and cybersecurity as well as its processes for identifying and designating essential services and critical entities, which were previously covered by two different directives.

This harmonization also constitutes a major step forward in improving the scope and nature of the EU CISR policy framework, with a new focus on enhancing security and resilience but still based on cross-sector and cross-border interdependencies. With the proposal for a resilience directive, the Commission's intent was to propose an all-hazards framework to increase the ability of critical entities "to prevent, resist, absorb and recover from disruptive incidents, no matter if they are caused by natural hazards, accidents, terrorism, insider threats, or public health emergencies" like the COVID-19 pandemic that has challenged the EU and the rest of the world since early 2020.[78] The proposal for a directive on resilience would expand the sectoral scope to include 10 sectors: energy, transport, banking, financial-market infrastructure, health, drinking water, wastewater, digital infrastructure, public administration, and space.[79]

To enhance the EU's collective CISR posture, the proposal includes specific measures to improve organization and oversight at the EU level and requirements for member states and critical entity operators to minimize vulnerabilities and ensure continued services.[80] Specifically, the proposal

76. European Commission, *Proposal for a Directive of the European Parliament and of the Council on Measures for a High Common Level of Cybersecurity across the Union, Repealing Directive (EU) 2016/1148* (Brussels: European Commission, 2020), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=72166.

77. European Commission, *Proposal for a Directive of the European Parliament and of the Council on the Resilience of Critical Entities* (Brussels: European Commission, 2020), 1–2, https://ec.europa.eu/home-affairs/system/files/2020-12/15122020_proposal_directive_resilience_critical_entities_com-2020-829_en.pdf.

78. European Commission, *Resilience of Critical Entities*, 14.

79. European Commission, *Resilience of Critical Entities*, 15.

80. European Commission, *Resilience of Critical Entities*, 10–12.

would require member states to have a strategy to ensure the resilience of critical entities, to conduct a national risk assessment, and to identify critical entities based on this risk assessment. The resilience proposal would also compel critical entities to conduct their own risk assessments, take appropriate technical and organizational measures to boost resilience, and report any disruptive incidents to their respective national authorities. At the EU level, the proposal would create oversight mechanisms, including advisory missions organized by the Commission, for critical entities providing services to or in at least one-third of member states. The Commission would offer different forms of support to member states and critical entities, such as an EU-level risk overview, best practices, methodologies, cross-border training activities, and exercises to evaluate the resilience of critical entities. Finally, the proposal would create the Critical Entities Resilience Group, comprised of experts, to facilitate regular cross-border cooperation to implement the directive.

The future of the EU CISR policy framework is at a critical juncture, as the European Parliament and the Council will likely issue these two directives by 2023 after discussing, negotiating, and updating the language put forth in the proposals. Initial indications suggest the process of negotiation will preserve most of the provisions envisaged in the proposals because the EU clearly needs to adopt these new measures to address the current and upcoming challenges pertaining to critical infrastructures, essential services, and critical entities.

EU's Future: Continuous Improvement and Adapting to New Threats

With the introduction of the NIS 2.0 and critical entity resilience directives in 2020—and their expected acceptance and implementation in the near future—the EU will have established an innovative, comprehensive, and inclusive framework that will prepare the member states to face the challenges to their critical infrastructures in the years to come. All the work conducted to improve security and resilience against physical and cyber threats will also pave the way for helping the EU and its member states prepare for and respond to hybrid threats, which pose a significant, complex, and concerning challenge for the EU. See chapters 2–4 for an overview of each of these types of threat.

Since 2016, the EU has been increasingly active in understanding the nature of hybrid threats and adopting measures to better prepare for and counter them. In particular, the Commission published two relevant proposals: (1) the *Joint Framework on Countering Hybrid Threats* in 2016 and (2) the 2018 *Joint Communication on Increasing Resilience and Bolstering*

Capabilities to Address Hybrid Threats.[81] More recently, in November 2020, the EU Commission's Joint Research Centre, together with the European Centre of Excellence for Countering Hybrid Threats, published a conceptual framework which describes the components of hybrid threats in terms of threat actors, their objectives, and their tools, as well as the domains they seek to compromise and the distinct phases of action in targeting these domains. This conceptual framework for hybrid threats constitutes another pillar of the recent innovations incorporated into the EU's CISR framework and already enables the member states to take initial steps to address the ever-evolving threat.

The timeline of the events described in this section on the EU provides an overall snapshot of the evolution and the current state of European CISR policies and practices. Consistent with its mentality of continuous improvement, the EU since 2004 has focused on adapting to address emerging threats, pursuing incremental changes, and allowing member states sufficient time to adopt and implement each improvement at their respective national levels. The EU approach has ensured a secure environment in which each critical infrastructure stakeholder has been able to discuss the most important matters, share information, become more familiar with the CISR efforts of other nations and within other sectors, get access to various best practices, and ask for bilateral or multilateral support as necessary. In this way, the EU's adaptive and comprehensive CISR framework puts the EU and its member states in a strong position to deter, prevent, reduce the consequences of, respond to, and recover from a broad array of threats to critical infrastructures and entities in the years to come.

81. European Commission, *Joint Framework on Countering Hybrid Threats* (Brussels: European Commission, 2016), https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:52018JC0016; and European Commission, *Increasing Resilience and Bolstering Capabilities to Address Hybrid Threats* (Brussels: European Commission, 2018), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018.

— 11 —

Information and Intelligence Sharing

Chris Anderson and Raymond Mey

Information and intelligence sharing are essential to the success of any critical infrastructure security and resilience (CISR) effort across the North Atlantic Treaty Organization because no single entity holds all the information necessary to fully understand:

- Threats—including intent, capability, and tactics, techniques, and procedures (TTPs)—from nation states and non-state actors, such as terrorists and criminals
- The vulnerabilities of infrastructure systems to those threats
- The far-reaching primary and cascading impacts of the degradation or loss of critical infrastructure, including dependencies and interdependencies across infrastructure facilities and sectors
- The most effective mitigation strategies against constantly evolving adversary TTPs

Given the complex and evolving information necessary to promote CISR, it is fundamentally important that the key infrastructure stakeholders endeavor to share information to fully understand comprehensive infrastructure risk so that they can determine the most efficient and effective means to mitigate these dangers. This process involves building trust, shared and practiced communications methods, and structured, multidimensional sharing.

Information-sharing Foundational Concepts

Effective information-sharing programs must be tailored to the legal, cultural, and social environments in which they operate. There is no "one-size-fits-all" approach to information sharing, but successful programs should consider how to incorporate the following foundational concepts of information sharing: value-added partnerships, the importance of trusted relationships, multidirectional sharing, and getting timely information to those who can act on it. This section will provide greater detail into each of these foundational concepts and conclude with a discussion on factors that can act as disincentives or barriers to sharing information, which successful information-sharing programs must take into account and strive to overcome.

Value-added Partnerships

Perhaps the most critical concept in building and maintaining an information-sharing program is that it provides information of value to all the participants. To make the necessary investments in personnel and processes, and be willing to share otherwise potentially sensitive information, members of an information-sharing group must be incentivized to continue participation, typically by ensuring that all participants have something to gain from the exchange of information. Information-sharing arrangements that are imbalanced risk losing participants who perceive no value in the investment if they are not receiving information of use. This point does not imply that mutual benefit is always balanced in the short term or even the long term, but without a clear reason for all participants to continue, programs will naturally decay over time.

In the most effective form of the value-added partnership, all participants develop and share information within their core area of expertise while benefiting from the expertise of others. Governments excel at collecting and maintaining large data sets of statistical or geospatial information. Military, law enforcement, and intelligence organizations often have the best (or the only) insight into adversary intentions and planning, while infrastructure operators maintain a deep understanding about the vulnerabilities and interdependencies that underpin systemic risk to the complex systems and processes they design and operate.

Intelligence, Information, and Data

While intelligence collection and information analysis lie beyond the scope of this chapter, there are two key concepts important to information-sharing frameworks. The first factor is the assessed credibility of information to be shared. Particularly when creating intelligence products about potential adversaries, information may be difficult to obtain and the adversary may seek to intentionally mislead such efforts. The second factor is the extent to which raw data must be analyzed and contextualized into information and intelligence to enable decision making and action.

These concepts are relevant to information sharing in critical ways. Robust information-sharing networks can greatly facilitate both credibility assessment and contextual analysis of raw data. But perhaps most importantly, accurate communication of the credibility and degree of analysis within a shared product is essential in building and maintaining the trust of recipients; there may be good reasons to share low-confidence or raw information, but it should always be communicated as such.

Importance of Trusted Relationships

Whether in a one-on-one conversation or a large, process-driven organization, the willingness to share information is related to the level of trust among the participants. Information that is of value is almost always of its nature sensitive to the holder of that information. Sharing may jeopardize the sources and methods by which it was developed, reveal strengths or weaknesses of a critically important facility or process, or allow information to be misused to build or erode competitive advantage. To overcome these sensitivities, participants in an information-sharing exchange must trust that recipients will use shared information for its intended purpose and will protect shared information from subsequent exposure or misuse.

Perhaps the most basic step in building levels of trust is to establish relationships ahead of time. Information-sharing circles should be established in steady state blue-sky conditions so that the participants and the organizations can get to know each other, group norms can be established, and basic processes can be developed, honed, and ingrained. Establishing these relationships early allows for time to identify and engage all the appropriate participants in each information-sharing group. With participants identified, mechanisms can be developed to share contact information so that information can be exchanged via e-mail, online forums, and through group meetings (in-person or virtual) as appropriate to the participants and the type

of information being shared. Regular engagement also ensures participants learn each other's professional jargon and facilitates rapid exchange of mission information in later situations when time may be of the essence. Finally, this blue-sky engagement allows for establishing processes, templates, and information-sharing rules that are critical to share information rapidly, effectively, and securely.

Multidirectional Sharing

To implement the two preceding concepts fully, an information-sharing regime must allow for, and even encourage, multidirectional information sharing. While any sharing arrangement may have times when the one-way flow of information is necessary—for example, governments may hold military, intelligence, or law enforcement information not in the public domain that is vital to infrastructure owners and operators—programs that are effective in the long term incorporate bidirectional information sharing. Information sharing cannot be limited to simply what government "pushes" to industry or programs that solely rely on industry to provide information—such as infrastructure vulnerabilities or suspicious activity related to their facilities or systems—to government partners without receiving information of value in return.

Multidirectional sharing goes even further than just two-way sharing between government and industry. At a foundational level, a government must work out how it shares information within itself (agency to agency) as well as across jurisdictions, from national to regional to local governments. Similarly, industry should examine available methods to share security information among various companies and across organizations within these companies. This multidirectional sharing often happens informally in cybersecurity circles as information security professionals leverage their personal networks and at the local levels among physical security professionals. As important as these methods are, they should not be relied upon as a substitute for more repeatable, formal, and survivable processes. Commercial providers of threat and risk information, particularly in the cyber realm, are increasingly vital in ensuring critical infrastructure operators are sufficiently informed of evolving threats.

To enable multidirectional sharing fully, participants must determine and follow norms regarding how shared information may be used, repackaged, combined with other elements, and disseminated without compromising any restrictions on sharing from the original provider of the information.

Timely Information to Those Who Can Act

Whether it is destined for the security guard at a facility gate in advance of a potential terrorist attack, a firewall administrator responsible for blocking potentially dangerous Internet traffic, or a first responder who can track down a suspicious person conducting possible surveillance of critical infrastructure, information is often extremely time-critical, particularly operational counterterrorism information. Given this time pressure, it is crucial to develop and improve processes constantly to speed the flow of information and get actionable information to those who will use it.

Speeding the flow of information requires careful analysis of every step in the intelligence- and information-sharing process. A common framework to evaluate a comprehensive intelligence process includes tasking, collection, processing, exploitation, and dissemination of information. While the first four steps in this process may also offer significant chances to decrease the time required to get information to recipients, this chapter primarily focuses on dissemination. Within the dissemination process, organizations should build procedures that push authority to release information to the lowest possible level. For highly time-sensitive information, this step may necessitate providing release authority to a watch operation center that runs 24 hours a day, seven days per week. Information shared on a recurring basis should have preestablished templates and formats known in advance so those who send the information and those who receive it can do so as rapidly and efficiently as possible. Within cybersecurity sharing efforts, these templates should also include machine-readable formats. Finally, information-sharing participants should develop common syntax and lexicon to facilitate rapid communication and understanding.

While getting information pushed out quickly is important, it is only part of the solution. The receiving organization needs to have procedures established to further route information to those within the group who will actually use it (if the initial recipients are not those who will act on the information). At the same time, transmitting organizations need to be careful not to overly classify or otherwise restrict how information can be further shared and disseminated. For example, passing classified or restricted information to a single cleared point of contact at a company may be of limited value if the information cannot then get to those who will ultimately take action. When sharing information with the private sector, this often means critical information should be at the lowest possible classification level or at a minimum have a "tear line" that conveys the most essential elements of information with the fewest restrictions. The blue-sky foundations

noted above enhance both the speed of information sharing and the standardized procedures that ensure information is shared in an actionable format to the right people at a critical infrastructure facility who can use it.

Information-sharing Disincentives

Any comprehensive discussion of information sharing must also address factors that deter or discourage information sharing. It is important to understand these natural disincentives and create programs and policies—including legislation or regulation where appropriate—that minimize their impact on the information-sharing environment.

The private sector has several potential disincentives to share information. In general, industry may be reluctant to share information voluntarily with the same government that is responsible for regulating the industry out of fear that sharing may invite enforcement action or encourage additional regulation. Antitrust or competitiveness law and regulation may preclude industry from sharing certain information with peers and competitors or even meeting with a group of competitors within the same sector. Industry may have concerns about voluntary disclosures that then become public and could create or exacerbate liability exposure by releasing nonpublic information later used in litigation against the releasing company. In some cases, laws or specific contract provisions may require that certain information is kept confidential, particularly customer proprietary network information or information related to health status. Industry may be reluctant to share perceived risks from other companies, such as supply-chain security concerns, due to defamation liability. For publicly traded companies, securities laws may limit public disclosure of material information about the company's operations or risks. Finally, industry may be concerned that proprietary information will fall into the hands of competitors, thereby damaging a company's position in the marketplace.

Government partners also have information-sharing impediments. Information held by intelligence agencies about nation-state or terrorist threats is often highly classified or otherwise tightly controlled, and the penalties for disclosing this type of information are severe because such information can in turn expose sources and methods used to gather the information. Federal and local law enforcement agencies are also concerned about compromising sources and information sensitive to ongoing investigations, leading to a reluctance to share timely information, and creating a gap between law enforcement and the private sector. To correct these deficiencies, efforts should focus on building trusting relationships and establishing

mechanisms such as "tear-line" products to share information and intelligence without compromising ongoing investigations, law enforcement effectiveness, and intelligence community tradecraft. Governments often have limits to naming a person or company as a threat without due process of law. Additional obstacles to government information sharing vary widely: (1) government agencies may be limited by law or regulation in sharing of certain types of information; (2) agencies may be reluctant to share information that may reflect negatively on their operations or capabilities; and (3) potential political or public relations fallout may make some government agencies hesitant to share potentially controversial information.

Information-sharing Subcategories

While these foundational concepts apply nearly universally, there are some specific considerations when dealing with certain types of information sharing. This section will describe the more narrowly defined information-sharing subtypes separately as they may offer distinct benefits to participants or may be amenable to specific types of information-sharing programs.

Cybersecurity

Defending against cyber threats offers one of the clearest examples of beneficial multidirectional information sharing. Within industry, many companies already share information with other companies within the same sector and even across sectors. Network defenders talk to counterparts at other companies to share observations about new threats and their associated TTPs. Private companies develop critical information about new vulnerabilities and mitigations for ongoing cyberattacks and often share this information openly in an effort to prove their bona fides and solidify their place in the market for cybersecurity prevention and mitigation services.

A core element of cyber information sharing is identification of the digital signatures of malicious activity and other ways to detect bad actors. Since cyber threats move at the literal speed of light, machine-to-machine sharing of this data is essential to the concept of getting timely information to those who can act. Information regarding common or newly discovered vulnerabilities is often critical, especially when accompanied by ways to eliminate or mitigate the vulnerability. When newly discovered malware is observed in the wild, critical infrastructure operators also benefit from, and are critical contributors to, shared information on corrective response actions that are effective against that specific cyber adversary or intrusion set.

Finally, robust information-sharing regimes will ensure private-sector companies—particularly in the communications and information technology sectors—are willing and able to share their situational awareness of the cyber environment with government in ways that enable all parties to observe, identify, correlate, and counter adversary activities in a rapid manner. Companies in the information-technology and communications sectors often have the best visibility into the cyber ecosystem over a large range of issues.

Physical Security

Organizations responsible for defending against terrorist attacks and other physical threats also benefit from robust information sharing. As adversary TTPs evolve, it is vital infrastructure defenders learn these new threats and then develop the most effective mitigations against them. Similar facility types often benefit from sharing best practices on defensive measures and benchmarking, a sort of comparison that can help identify where a facility's physical security posture is not as strong as its peers. Ensuring a baseline defensive level helps facilities from being singled out as the easiest or softest target. Since many critical infrastructure operators constantly observe the areas around their facilities, their vigilant observation and reporting may help identify suspicious activity or pre-attack planning. Reporting suspicious activities (such as photography of perimeter defenses or unmanned aerial vehicle overflights) may indicate ongoing surveillance or planning activity. Regular information sharing may help identify an adversary team operating in a region and surveilling multiple facilities to select a target. Military, law enforcement, or intelligence services may have the best information available regarding the TTPs of potential adversaries through battlefield observation, group infiltration, or intercepts of planning activity. Sharing these TTPs with critical infrastructure owners and operators will enable them to understand the potential consequences of new attack methods and develop appropriate countermeasures.

Risk Analysis and Mitigation

Government and industry partnership is essential to understanding risks to critical infrastructure and informing both public and private investments to mitigate these risks. Without industry insights, it can be extremely difficult for government planners to understand the vulnerabilities of highly complex infrastructure facilities or the consequences of various attacks on them, whether in the physical or cyber domains. Even more challenging is gaining a full appreciation of the cross-sector dependencies and interdependencies—

the critical connections that can increase or mitigate risks. See chapter 12 for an overview of the nature of interdependencies and of using a system of system approach to enhance resilience. Good information sharing can greatly contribute to mutual appreciation of common vulnerabilities, which can serve to identify areas for industry action, government research and development, or other mitigation measures. Recent events have placed a spotlight on supply-chain risks for both physical hardware and critical software. Shared government and industry analysis and data sets can help uncover potential risks within both physical and cyber supply chains from complex cross-sector relationships, physical or virtual chokepoints, and concentration of critical industrial sources in a single supplier, country, or region.

Information-sharing Regimes and Programs

This section describes a range of existing information-sharing regimes currently in operation in the United States and Europe. Many of these programs are designed to maximize the benefits or address the disincentives described earlier in this chapter. While some of these programs overlap areas of responsibility—because information sharing can be highly dependent on relationship building—it can be beneficial to have a range of programs which provide "defense in depth" and help ensure broad dissemination of crucial information.

In the United States, the Department of Homeland Security (DHS) is charged with coordinating CISR efforts.[1] Working with other federal departments, DHS operates a range of programs and initiatives intended to build a robust public-private voluntary partnership founded on information sharing. This section will briefly introduce four types of these programs, the first of which is the Critical Infrastructure Partnership Advisory Council (CIPAC). CIPAC is a legal framework that allows interaction between the government and critical infrastructure owners and operators, enables industry representatives to develop and share consensus recommendations with the federal government, and facilitates government and industry discussions on security and resilience topics in nonpublic engagements.[2] In countries where antitrust restrictions or disclosure requirements limit the government's ability to meet nonpublicly with private companies individually or as a group,

1. Homeland Security Act of 2002, H.R. Rep. No. 107-609, pt. 1 (2002).

2. "Critical Infrastructure Partnership Advisory Council," Cybersecurity and Infrastructure Security Agency (website), n.d., accessed September 29, 2021, https://www.cisa.gov/critical-infrastructure-partnership-advisory-council.

a framework such as CIPAC may facilitate forthright government-industry discussions necessary to solve security and resilience challenges.

A second information-sharing program is the Protected Critical Infrastructure Information (PCII) program. One major obstacle in government-industry information sharing is the general reluctance among private-sector companies to provide the government information that could later be used against them in regulatory or liability actions. To overcome this obstacle, the United States adopted the PCII legal framework. The PCII program protects information voluntarily provided by industry against use for regulatory purposes and disclosure under some "Freedom of Information" government obligations to disclose information to the public.[3]

Given that information held by governments regarding threats to critical infrastructure may well be classified, it is important to work through ways to share this information with industry to help prevent attacks or mitigate potential consequences. As noted above, the most important strategy to overcome this challenge is to develop processes to create tear-line products that can be shared at an unclassified level. This step is particularly important with any recommended protective measures so these measures can be shared within and across private-sector companies freely. In addition to using tear-line products, DHS has also adopted the Private Sector Clearance Program, which grants a limited number of clearances to private-sector personnel who are responsible for securing critical infrastructure facilities so classification issues do not hinder the government from sharing vital information when necessary. Since some information simply cannot be declassified, some Allies and partner nations may find it useful to pursue programs similar to this one.

While the previous programs apply to both cyber and physical information sharing, a fourth type of program—cybersecurity information sharing—has some unique characteristics that require specialized approaches to information sharing. Recognizing this need led the DHS to develop the Cyber Information Sharing and Collaboration Program, which enables timely exchange of unclassified threat and vulnerability information through trusted public-private partnerships.[4] With cybersecurity information sharing, defining a common language and syntax is important, particularly if information sharing is to be done from machine to machine—

3. "Protected Critical Infrastructure Information (PCII) Program," Cybersecurity and Infrastructure Security Agency (website), n.d., accessed September 29, 2021, https://www.cisa.gov/pcii-program.

4. "Cyber Information Sharing and Collaboration Program (CISCP)," Cybersecurity and Infrastructure Security Agency (website), n.d., accessed September 29, 2021, https://www.cisa.gov/ciscp.

a critical capability given the speed at which new cyber threats proliferate. One such system is the Automated Indicator Sharing, which facilitates the exchange of machine-readable cyber threat indicators and recommended protective measures.[5] This sharing capability uses open standards, such as the Structured Threat Information Expression for cyber threat indicators and defensive measures information, and the Trusted Automated Exchange of Indicator Information for machine-to-machine communications.[6]

Similarly, the US Federal Bureau of Investigation (FBI) coordinates several additional information-sharing programs, including the three examples discussed here. The first program is the Joint Terrorism Task Force (JTTF) concept. When the FBI's New York Division established the first JTTF in 1980, the task force—comprised of members from the New York Police Department and the FBI—enabled the sharing of vital information pertaining to bank robberies conducted by a terrorist group in New York. Given its success in addressing the need to share information and intelligence in support of these investigations, the JTTF concept proved valid for future use as well.

In 1998, the FBI formed a JTTF to begin security preparations for the 2002 Winter Olympic Games hosted in Salt Lake City, Utah. From 1998–2002, this task force focused its investigation on a right-wing, anti-Semitic group known as the Sons of Aryan Culture, which was operating a criminal enterprise throughout the Salt Lake City metropolitan area. The JTTF's investigative efforts revealed the gang was developing plans to attack the Olympics, but over the course of four years, the JTTF was effective at dismantling this group and interrupting its plans to target the games. As part of the JTTF effort, over 100 leaders in the public-safety community received top-secret security clearances to facilitate the legal sharing of sensitive information and intelligence. This effort proved to be highly effective in countering terrorism threats to the Olympics and establishing an efficient and unified public-safety effort in support of this major event. Perhaps the most significant JTTF effort occurred after the 9/11 terrorist attacks, when the JTTF included over 100 different groups and incorporated 5,000 local, state, and federal officers.

Notably, the scope of JTTF membership after 9/11 expanded and now can incorporate public-safety organizations within a given jurisdiction.

5. "Automated Indicator Sharing (AIS)," Cybersecurity and Infrastructure Security Agency (website), n.d., https://www.cisa.gov/ais.

6. "Sharing Threat Intelligence Just Got a Lot Easier!," OASIS Cyber Threat Intelligence Technical Committee (website), n.d., accessed September 29, 2021, https://oasis-open.github.io/cti-documentation/.

The critical component of individual JTTF success has been the proven willingness to share intelligence and information among participating agencies. Members of the JTTF undergo extensive security background checks before being allowed to join a JTTF. Although costly and time consuming, the background checks and security clearances required of each JTTF member allow for coordinated interagency investigations and sharing of vital sensitive information and intelligence. JTTFs have a tremendously effective law enforcement mechanism for sharing information and sensitive intelligence and have greatly contributed to counterterrorism success in the United States since 9/11.

The second significant information-sharing program under the FBI is a geographically focused network for critical infrastructure information sharing known as InfraGard. The network provides a vehicle for public-private collaboration intended to expedite the timely exchange of information and promote mutual learning opportunities relevant to critical infrastructure defense. One of the program's key strengths is the connection of local InfraGard chapters to the local FBI field offices. This geographic connection allows for blue-sky relationship building and the establishment of relationships based on mutual trust. During steady-state operations, InfraGard focuses on training and information sharing through webinars, unclassified threat briefings, and document sharing through a protected online portal. InfraGard operates around 80 chapters across most of the major US metropolitan areas and engages leaders, often at the facility level, who are on the "front lines" of infrastructure protection.

A third information-sharing program run by the FBI is the Domestic Security Alliance Council (DSAC). In contrast to InfraGard's focus on building trusted relationships with operators at a local level, the DSAC provides an executive-level mechanism to enhance communication and promote timely and effective exchange of security and intelligence information between the federal government and the private sector.[7] To accomplish this objective, the DSAC brings together senior government leadership from the FBI and DHS along with private-sector executives. The DSAC, formed in 2005, was modeled after the US Department of State's Overseas Security Advisory Council, which works with the private sector to help ensure US private-sector companies operating overseas are aware of threats to international operations.

7. "About DSAC," Domestic Security Alliance Council (website), n.d., accessed December 21, 2021, https://www.dsac.gov/about.

A range of programs in the European Union focus on sharing critical infrastructure security and resilience information for both cyber and physical security. The European Commission established one of the more recent programs on June 23, 2021, when it announced the establishment of the Joint Cyber Unit (JCU) to foster cooperation between the cyber communities across the EU institutions, agencies, bodies, and authorities in the member states. Specifically, the JCU aims to enhance cooperation and information sharing among the various resilience, law enforcement, defense, diplomacy, and private-sector entities—collectively, the users and providers of cybersecurity solutions and services—in response to the increase in major cyber incidents impacting citizens and businesses across the EU. To achieve its goals of facilitating an EU coordinated response, improving situational awareness, and guaranteeing joint preparedness, the JCU will develop capabilities over four planned stages of growth and is expected to become fully operational by mid-2023.[8]

One information-sharing program common in the United States and Europe is an Information Sharing and Analysis Center (ISAC), which is a critical infrastructure sector-specific organization to share information about threats and vulnerabilities. The EU Agency for Cybersecurity is a proponent of ISACs in Europe, and its cooperative models guide provides a comprehensive description of ISAC operations in Europe, including country and sector-specific models as well as internationally focused ISACs.[9] In the United States, many infrastructure sectors and subsectors self-organize into an ISAC that meets the needs of the member companies. Some ISACs provide fee-for-service, while others closely coordinate and leverage government operations centers for information sharing. Some ISACs, such as the US Communications ISAC, are broad in scope and cover an entire sector, while others are highly focused, like the Downstream Natural Gas ISAC. ISACs can also coordinate to share information from ISAC to ISAC, enabling cross-sector information exchange to promote greater infrastructure security and resilience.

Information-sharing programs are by no means limited to those directed top-down from government agencies. Many of the most effective arrangements grow organically to meet the needs of participants linked by a common geography or similar operational environments. While federal

8. "Shaping Europe's Digital Future: Joint Cyber Unit," European Commission (website), n.d., accessed September 28, 2021, https://digital-strategy.ec.europa.eu/en/policies/joint-cyber-unit.

9. "Information Sharing and Analysis Center (ISACs)—Cooperative Models," European Union Agency for Cybersecurity (website), n.d., February 4, 2018, https://www.enisa.europa.eu/publications/information-sharing-and-analysis-center-isacs-cooperative-models.

government programs described earlier can be important to the overall information-sharing success, many programs have evolved without leadership, support, or sometimes even involvement of a federal government. A variety of local, state, national, and international groups advance information-sharing capabilities for their members through operational support or simply defining templates, standards, and processes that facilitate information sharing.

For example, FIRST—an international partnership focused on bringing together cyber incident response and security teams—has documented a simple "traffic light protocol" that specifies how a given report or piece of information can be further shared or restricted.[10] This system provides some degree of security to the sharer of information, but also clarity to the recipients about whether and with whom the information can be shared. In the United States, state and local fusion centers promote law-enforcement information sharing across federal, state, local, and private sector communities within a given state or region. National and international professional associations, such as the Institute of Electrical and Electronics Engineers and the American Society of Civil Engineers, help share information about security and resilience of their respective infrastructures. Finally, industry groups often unite to share information on key topics of interest in critical infrastructure resilience, such as the Council to Secure the Digital Economy's International Botnet and IoT Security Guide or the American Waterworks Association's sharing of information vital to securing the water sector.[11]

Case Studies: Information Sharing in Action

This section briefly outlines several real-world examples of information sharing. Each case study will highlight one or moreof the concepts and programs highlighted in the previous section.

Cyber Health Working Group: Public-Private Information Sharing

The Cyber Health Working Group (CHWG) was established in 2015 by the FBI Washington Field Office, the InfraGard National Members Alliance, the InfraGard National Capital Region chapter, and the Executive Partnership for Integrated Collaboration, a 501(c)(3) nonprofit in Charlotte, North Carolina. Using a simple Listserv technology, the CHWG provides

10. "Traffic Light Protocol," FIRST (website), n.d., accessed September 29, 2021, https://www.first.org/tlp/.

11. "International Botnet and IoT Security Guide," Council to Secure the Digital Economy (website), n.d., accessed September 29, 2021, https://csde.org/projects/international-anti-botnet-guide/.

a two-way, real-time information-sharing platform for cyber practitioners in the health care and public-health sectors. Members share best practices, emerging threats and trends, and indicators of compromise. Since the group's inception, membership has grown from approximately 200 at the beginning to more than 1,300 members.

Participants are InfraGard members at the intersection of the information-security and health-care sectors who use a Listserv and online portal to share information about risk, compliance, emerging threats, and best practices in cybersecurity. As InfraGard is a public-private collaboration program, the CHWG expressly includes members from both the private and public sectors. The CHWG's steering committee, comprised of CHWG private-sector participants who represent multiple subsectors of the health-care sector, provide basic governance for the group and set the mission, scope, and parameters for membership. Among its various private and public-sector members, CHWG's mission is to develop, foster, and facilitate a community comprised of cybersecurity-focused professionals working in the health-care sector, those who have responsibilities (such as physical facility cybersecurity) and those who can share real-time information about risk, governance, threats, indicators, trends, and best practices. Membership of the CHWG, as established by the steering committee, is open to any individuals who meet the following criteria:

- Is a current InfraGard member or has an application pending with the local InfraGard chapter.
- Works in cybersecurity for an organization in the health-care sector or handles cybersecurity for a physical facility.
- Is at least partially responsible for cyber threat, risk, governance, or compliance issues.
- Can access and potentially share threat information and/or cybersecurity best practices.
- Does not have a role that is primarily focused on business development or sales of products and services and will not use the group to market or promote products and services.

The CHWG, according to the FBI's Washington Field Office, has aided the FBI in its mission by predicating cases, enhancing ongoing investigations, identifying new victims of cyber intrusions, and contributing to several intelligence products. For example, during a recent cyberattack on a major health-care provider, the CHWG identified the attack before the FBI's 24/7 cyber operations watch center

and group members in multiple states shared substantive information and identified other information-sharing networks that had relevant information, including indicators of compromise. Following the success of the CHWG, additional working groups using the same model have been established for the commercial facilities sector (cybersecurity), the finance sector (cybersecurity), and the data center sector (physical security).

If You See Something, Say Something®

Originally implemented and trademarked by the New York Metropolitan Transportation Authority, If You See Something, Say Something® is a US national anti-terrorism outreach program licensed to the DHS. In partnership with the Department of Justice's Nationwide Suspicious Activity Reporting Initiative, the DHS officially launched this campaign in July 2010 with the goal of raising awareness of the indicators of terrorism and terrorism-related crime and training law enforcement at the state and local levels to recognize these types of behavior and indicators.[12] The Nationwide Suspicious Activity Reporting Initiative developed a standard process for documenting and analyzing the observations the campaign generates, and it routinely shares these reports with relevant FBI-led JTTFs for investigation and with state fusion centers for analysis.

Since its beginning, the If You See Something, Say Something campaign has been expanded to include states, counties, cities, and transportation entities (such as airports and mass transit, major entertainment venues and sports events, colleges and universities, private-sector businesses, and media outlets). Program partners promote the campaign as part of their larger safety and security plan using internal advertisements for employees (for example, in employee common areas, breakrooms, restrooms, and via e-mail) in addition to public-facing areas. Organizations that join the campaign have access to public-service announcements, educational materials, signage, and relevant and enduring social media content.

The campaign has been responsible for disrupting several terrorist plots and saving lives.[13] One example is the case of Alexander Ciccolo of Adams,

12. "If You See Something, Say Something," Department of Homeland Security (website), n.d., accessed September 28, 2021, https://www.dhs.gov/see-something-say-something.

13. Bridget Johnson, "5 Times Terrorists Were Thwarted by 'If You See Something, Say Something,'" Homeland Security Today (website), September 25, 2018, https://www.hstoday.us/subject-matter-areas/counterterrorism/5-times-terrorists-were-thwarted-by-if-you-see-something-say-something/.

Massachusetts. The son of a Boston police captain, Ciccolo went by the name Ali Al Amriki and developed an admiration for the Islamic State (Da'esh) terrorist group. Based on a tip from Ciccolo's father, the FBI initiated an investigation during which Ciccolo was recorded talking about plans to engage in terrorist activity, inspired by the Islamic State (Da'esh), that involved filling pressure cookers with black powder, nails, and ball bearings. Following a July 2015 sting operation, Ciccolo was arrested and subsequently pleaded guilty to several charges, including attempting to provide material support to a foreign terrorist organization. In September 2018, he was sentenced to 20 years in prison and a lifetime of supervised release.[14]

Attack on the US Capitol: An Information-sharing Failure?

On January 6, 2021, rioters attacked the US Capitol building in an attempt to disrupt a joint session of Congress, during which members of Congress were to count the electoral votes for president and vice president of the United States, and then announce the official results of the 2020 election. The attackers breached the Capitol building, vandalized and stole property, ransacked offices, attacked members of law enforcement, and threatened the safety and lives of elected US leaders. Capitol police officers, along with federal, state, and local law enforcement partners, reestablished control of the building, and the President of the Senate, Vice President Mike Pence, announced Joseph R. Biden Jr. and Kamala Harris as the president-elect and vice president-elect of the United States in the early morning hours of January 7. Tragically, seven individuals, including three law enforcement officers, lost their lives.

The US Senate investigated the security, planning, and response failures and critical breakdowns at various federal agencies, particularly the FBI, the DHS, and Department of Defense. The Senate investigation documented a number of failures that set the stage for the breach of the Capitol. Of interest here, the report specifically highlighted two information-sharing failures as being significant to the overall failure to protect the Capitol.

First, the Senate report found the federal intelligence community—led by the FBI and the DHS—failed to issue a threat-assessment warning

14. "Massachusetts Man Inspired by ISIS Sentenced for Plotting to Engage in Terrorist Activity," Department of Justice Office of Public Affairs (website), September 5, 2018, https://www.justice.gov/opa/pr/massachusetts-man-inspired-isis-sentenced-plotting-engage-terrorist-activity.

of potential violence targeting the Capitol on January 6.[15] The report documents how enforcement entities, including US Capitol Police (USCP) law, rely on the intelligence community to assess and communicate homeland security threats. While the FBI and DHS disseminated multiple written documents detailing the potential for increased violent-extremist activity and targeting of law enforcement personnel and government facilities and employees at other lawful protests throughout 2020, no such reports were released regarding the threats on January 6. Although online calls for violence at the Capitol were observed and assessed by the agencies, FBI and DHS officials stated that observed activity was possibly constitutionally protected free speech instead of actionable, credible threats of violence. In their testimony, officials from both the FBI and the DHS acknowledged the need to improve handling and dissemination of threat information from social media and online message boards. This first Senate report finding shows how disincentives—in this case a combination of unclear legal guidelines and concern for political second-guessing—can limit a government agency's willingness to share information.

Second, the Senate report found the intelligence components within the USCP, which are responsible for generating their own intelligence analysis, did not convey the full scope of threat information they possessed.[16] The report documented that—despite the lack of FBI and DHS Intelligence and threat information—USCP's lead intelligence component, the Intelligence and Interagency Coordination Division (IICD), was aware of the potential for violence in the days and weeks ahead of January 6, based on multiple information sources regarding the large crowds expected to gather in Washington on January 6 and specific threats of violence focused on the joint session of Congress and the Capitol building complex. Even though IICD had the necessary information internally, it failed to incorporate this information fully into internal assessments about January 6 and the joint session. As a result, USCP officers and other law-enforcement partners did not have this critical information regarding threats of violence. This finding shows how damaging it can be when vital information is not shared with those who need it to take protective actions to mitigate risks and threats, and secure critical infrastructure.

15. Committee on Homeland Security and Government Affairs and Committee on Rules and Administration, *Examining the U.S. Capitol Attack: A Review of the Security, Planning, and Response Failures on January 6* (Washington, DC: US Senate, 2021), 1, https://www.rules.senate.gov/imo/media/doc/Jan%206%20HSGAC%20Rules%20Report.pdf.

16. Committee on Homeland Security, *Examining U.S. Capitol Attack*, 2.

Further, the Senate report asserts USCP's preparations for the joint session suffered because of the decentralized nature of its intelligence components and lack of sharing across those elements.

> On January 5, an employee in a separate USCP intelligence-related component received information from the FBI's Norfolk Field Office regarding online discussions of violence directed at Congress, including that protestors were coming to Congress "prepared for war." This report, which was similar to other information received by IICD and could have served as corroborating information, was not distributed to IICD or USCP leadership before January 6.[17]

This is an unfortunate example of the importance of multidirectional information sharing; had different components within the same structure been more effective at lateral information sharing, the outcome of the events on January 6 might have been very different.

National Terrorism Advisory System

Following the 9/11 terrorist attacks on the United States, the government determined a need for a standard way to share information regarding terrorism threats with the general US population. The first of these programs, the Homeland Security Advisory System (HSAS), was established on March 12, 2002.[18] This program was intended to be a simple means to communicate threat levels and, as such, provided five basic color-coded threat levels (see figure 11-1).[19] In addition to the threat levels, HSAS also provided recommended protective measures for federal agencies but, notably, not for state and local governments, the private sector, or the general public. The simple, preestablished color-coded system leveraged one of the key principals for information sharing, which is to establish templates and definitions in advance to facilitate quicker sharing of complex information.

However well-intentioned, HSAS fell victim to several pitfalls common to information-sharing programs. One of the primary issues with HSAS

17. Committee on Homeland Security, *Examining U.S. Capitol Attack*, 2.

18. Congressional Research Service (CRS), *Homeland Security Advisory System: Possible Issues for Congressional Oversight* (Washington, DC: CRS, January 29, 2008), 1, https://crsreports.congress.gov/product/pdf/RL/RL32023.

19. "Gov. Ridge Announces Homeland Security Advisory System," White House Office of the Press Secretary (website), March 12, 2002, https://georgewbush-whitehouse.archives.gov/news/releases/2002/03/20020312-1.html.

was that its warnings were vague and, for most users, came without specific actionable protective measures. As discussed earlier, information-sharing programs function best when they place timely, actionable information in the hands of those able to implement protective measures. While HSAS changes were shared through a range of mechanisms, none of these were wholly satisfactory, and many users reported to congressional oversight committees their primary source for HSAS notifications was open-source news media. Raising HSAS levels often created significant costs for federal agencies, state and local responders, and the private sector as they implemented tighter protective measures. Those responsible for determining HSAS levels were therefore reluctant to raise the levels, while, at the same time, they did not want to lower the threat level only to be surprised by a hard-to-detect terrorist attack and be accused of letting down their guard at precisely the wrong moment. The net result was HSAS threat levels did not change very often; the guarded and low levels were never used, and the highest level (severe), was used only for three days but only specifically for inbound flights from the United Kingdom immediately following an attack there.[20]

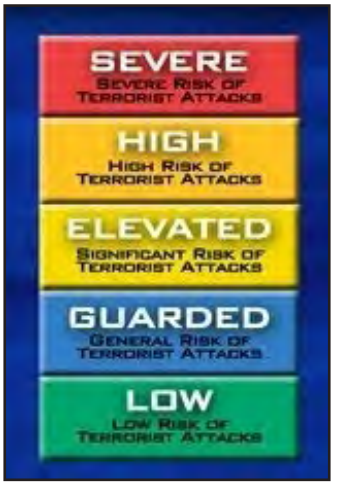


Figure 11-1. Homeland Security advisory system

In January 2011, DHS retired HSAS and replaced it with a new system, the National Terrorism Advisory System (NTAS). The new system eliminated the color-coded levels in favor of two specific products: elevated alerts that warned of credible threats and imminent alerts that warned of a near-term credible threat that suggested a need for immediate protective actions. Both types of alerts had a high threshold to prevent the overuse of the system,

20. "Chronology of Changes to the Homeland Security Advisory System," Department of Homeland Security (website), September 13, 2017, https://www.dhs.gov/homeland-security-advisory-system.

which might cause recipients not to take alerts seriously. This approach had a downside, however, as it precluded the sharing of lower-level information that did not meet these high thresholds, but could still be of use in defending against threats. To fill this gap, DHS later added a new category of advisory, known as bulletins, to share information regarding "terrorism trends, events, and potential threats in those situations where additional precautions may be warranted, but where the circumstances do not indicate a threat against the United States of sufficient credibility, or specificity and credibility" to justify issuing an alert.[21] While NTAS has offered some substantive improvements, it has not entirely overcome the vagueness, lack of specific protective measures, and disincentives to use that hindered its predecessor.[22]

National Special Security Events and Special Event Assessment Rating

Since major events—such as presidential inaugurations, the Olympic Games, the World Cup, and the Super Bowl—attract millions of spectators and worldwide attention or bring critical leaders together in one place, they demand special security considerations to protect against terrorist attacks adequately. In fact, several of these types of events have been the past targets of terrorist plots and attacks. In preparation for these events, host nations may designate such events as matters of national security and allow for special national funding, authorities, and resources—including information-sharing efforts—to assist in planning and operational requirements. In the United States, the federal government may designate certain events as national special-security events (when the president, foreign heads-of-state, or other dignitaries under Secret Service protection will be present) or assign a special event activity rating (when large crowds or symbolic events may elevate the threat and consequence of a terrorist attack).

As previously stated, trust is a critical element that must be developed for agencies and organizations to be motivated to share information and intelligence. Designating these major events provides the resources and focus necessary to allow agencies to interact, share jurisdictional planning and operational responsibilities, and develop trusting relationships through integrated, coordinated security and counterterrorism efforts.

21. "NTAS Frequently Asked Questions," Department of Homeland Security (website), n.d., accessed June 15, 2021, https://www.dhs.gov/ntas-frequently-asked-questions.

22. Matthew Wein, "Back to Threat-Level Orange and the Need to Update the National Terrorism Advisory System," *Lawfare* (blog), December 6, 2016, https://www.lawfareblog.com/back-threat-level-orange-and-need-update-national-terrorism-advisory-system.

CISR is key to the success of these major events because the security of the participants and successful execution of the event always rely on core infrastructure services, such as reliable power, communications, transportation, and water. Building trust and coordinating operations between different critical infrastructure sectors require a close working relationship with the private sector, law enforcement, and security components. The inclusion of critical infrastructure owners and operators in the planning and operational coordination is no small feat, but such efforts develop the foundation for trust for all involved, encourage the sharing of information and intelligence, and ultimately assist in securing the critical infrastructure and, as a result, the event itself.

Summary and Actions for Consideration

Information-sharing programs will always present operational and policy challenges because so many of their elements hold internal tension and must therefore be constantly balanced. Sharing too much can result in information fatigue, but providing too little can mean critical knowledge does not get where it can be put to use. Similarly, shared information must find the happy medium between (1) reporting that is too exhaustive and therefore obscures the critical points in irrelevant data, and (2) superficial reporting that lacks sufficient detail to inspire confidence and allow for rapid and appropriate decision making.

As discussed throughout the chapter and drawn out in detail in the case studies, several keys to building effective information-sharing programs are listed below.

- Build trusted relationships with the right partners. The importance of trusted relationships to effective operational information sharing cannot be overstated. This trust often can be cultivated through regular contact during blue-sky or steady-state conditions. Built into this consideration is careful identification, recruitment, and cultivation of the critical members of an information-sharing community, namely, those who have critical pieces of information and those who can take action based on the information to be shared.
- Establish the sharing norms within a community. Community norms include: (1) guidelines or rules for use and further sharing of information received through the network;

(2) a common lexicon and terminology, including thresholds and criteria for when and how to share information; and (3) templates and formats for information-sharing products. These templates could be as specific as machine-readable cybersecurity message formats like STIX/TAXII or human-readable report formats that include critical information where recipients expect to see it.

- Evaluate the incentives and disincentives to share information and tailor programs to maximize benefits and minimize drawbacks for all participants. It is not sufficient to encourage information sharing and expect it to occur suddenly and naturally. Establishing an effective program requires understanding the different perspectives on incentives and disincentives different stakeholders hold and then devising specific programs to extend the advantages and limit the disadvantages.
- Consider the medium in which sharing activity will take place. Information-sharing mechanisms often involve tradeoffs, particularly between security and ease-of-use. Depending on the needs of a community, sharing could take place in person only; through multifactor authenticated online forums; through e-mails, phone calls, and texts; or by leveraging open-source news and broadcast media.
- Constantly reassess effectiveness. No information-sharing program perfectly meets the needs of its community, and even so, community needs evolve over time. Regular reassessment allows the chance to ensure sharing processes are as rapid as possible and that information is actually getting to those who can use it most.

To secure NATO critical infrastructure effectively against the full range of threats, including terrorism, decisionmakers in military, law enforcement, and owner-operator communities must collaborate to share the different puzzle pieces of information each may hold. Only through trusted sharing relationships in an established, secure environment will a comprehensive risk picture materialize. This process will allow risk-informed decisions about protective measures and resilience investments to drive more secure critical infrastructure.

— 12 —

Critical Infrastructure Interdependency Modeling and Analysis: Enhancing Resilience Management Strategies

Duane Verner

North Atlantic Treaty Organization member states and partner nations face significant challenges in formulating and implementing effective strategies to address the risks posed by natural, man-made (to include terrorist), and socio-technological hazards. Extreme weather events have the potential to cause massive damage to multiple sectors of infrastructure following a single storm.[1] Disruptive effects have resulted from the mismanagement and lack of investments in infrastructure supporting lifeline community needs.[2] Perhaps the most pressing of these challenges, climate change, poses significant threats to coastal communities as well as shifting environmental conditions across the world—such as temperature and precipitation—for which current infrastructure design and performance has not

Acknowledgments: The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a US Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. Frédéric Petit, Joshua David Bergerson, Matthew Berry, Lawrence Paul Lewis, and Duane Verner contributed to this chapter. Duane Verner served as the corresponding author for the duration of the CISR handbook project. The authors are particularly grateful to the late Dr. James Peerenboom. Thank you, Jim, for setting the stage for homeland security research. We hope to be able to follow your example and contribute to the advancement of critical infrastructure interdependencies assessments.

1. Ricardo Rosselló, *Build Back Better Puerto Rico: Request for Federal Assistance for Disaster Recovery* (San Juan: Government of Puerto Rico, 2017), 2, https://media.noticel.com/o2com-noti-media-us-east-1/document_dev/2017/11/13/Build%20back%20better%20Puerto%20Rico_1510595877623_9313474_ver1.0.pdf.

2. Flint Water Advisory Task Force, *Final Report* (Lansing, MI: Office of Governor, March 21, 2016), 1–2, https://www.michigan.gov/documents/snyder/FWATF_FINAL_REPORT_21March2016_517805_7.pdf.

been adapted.[3] The complexity involved in assessing these risks and managing the potential disruptive impacts of these events to critical infrastructure increases when considering the inherent interdependencies that exist between infrastructure assets and systems. Critical infrastructure assets operate in concert with one another. Catastrophic events can therefore cascade across these interconnected systems and hamper the ability of critical infrastructure operators to remain operational.

Policy guidance across disciplines, from homeland security principles to business continuity standards, reflects the importance of considering interdependencies to enhance the security and the resilience of critical infrastructure systems.[4] The majority of these documents, however, do not comprehensively define methodologies to understand complex interactions among critical infrastructure and to support resilience management strategies. In recent years, research and development in modeling critical infrastructure interdependencies has developed considerably. These approaches generally use systems engineering and safety engineering techniques.[5] Applying these techniques to critical infrastructure systems, however, is challenging due to the difficulties associated with obtaining all the data necessary to run combined simulation models of multiple infrastructure systems.[6]

To address these challenges, a system of systems service-oriented approach can help to establish the appropriate scope of an interdependency analysis and prioritize the specific assets and/or subsystems for which resilience-related information should be collected. Using this approach, analyses consider the high-level context—socioeconomic fragility, for example—of a given region, the characteristics of critical infrastructure assets providing resources to the region, and the capabilities of both the critical infrastructure and the

3. Megan Clifford et al., "Closing the Gap between Climate Science and Critical Infrastructure Adaptation," George Mason University Center for Infrastructure Protection and Homeland Security (website), August 6, 2015, https://cip.gmu.edu/2015/08/06/closing-the-gap-between-climate-science-and-critical-infrastructure-adaptation/.

4. For additional information, see White House, *Presidential Policy Directive (PPD)-21—Critical Infrastructure Security and Resilience* (Washington, DC: Office of the Press Secretary, February 12, 2013), https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil; and ASIS International, *Organizational Resilience: Security, Preparedness, and Continuity Management Systems—Requirements with Guidance for Use* (Alexandria, VA: ASIS International, 2009), https://www.ndsu.edu/fileadmin/emgt/ASIS_SPC.1-2009_Item_No._1842.pdf.

5. For additional information, see Min Ouyang, "Review on Modeling and Simulation of Interdependent Critical Infrastructure Systems," *Reliability Engineering & System Safety* 121 (January 2014): 43–60, https://doi.org/10.1016/j.ress.2013.06.040.

6. For additional informatoin, see Megan Clifford and Charles Macal, *Advancing Infrastructure Dependency and Interdependency Modeling* (Argonne, IL: Argonne National Laboratory, 2016), https://anl.app.box.com/s/3t7mnesdajzl708xy9xo4vczj2qv1wom.

regional emergency management. By specifically addressing stakeholders' requirements and building a collaborative environment, a flexible analysis framework can enable the development of a comprehensive and interactive resilience assessment of critical infrastructure interdependencies. This framework can also serve to integrate multiple areas of expertise—such as engineering, social sciences, business continuity, and emergency management—in a combination of top-down (system-level) and bottom-up (asset-level) approaches.

This chapter explores the potential of this analysis framework approach in the following five sections. The first section discusses the importance of considering interdependencies to improve the security and resilience of critical infrastructure systems. The next section provides a general overview of the main characteristics and dimensions of critical infrastructure interdependencies. The third section summarizes general approaches to model and assess critical infrastructure. The next section proposes a flexible critical infrastructure analysis framework to inform the development of resilience management strategies. The final section identifies elements to operationalize the consideration of critical infrastructure interdependencies for managing critical infrastructure systems. NATO Allies and partners can use the proposed analysis framework and complementary considerations for operationalizing its results to reduce the risks posed to critical infrastructure and to foster greater resilience through cross-sector collaboration.

Risk, Resilience, and Interdependencies

Analysts frequently identify risk and resilience as essential aspects in ensuring the continued safe operation of infrastructure in a cost-effective manner. Although some variation exists in the scholarship as to its definition, *risk* is generally the combination of the magnitude of a consequence—such as a loss following a disruption event—and the probability or likelihood of the consequence occurring.[7] Stanley Kaplan and B. John Garrick first identified three key questions that could drive risk assessments. What are the problems that could occur in the system? What is the probability that those problems could occur? What would be the consequences of those occurrences?[8] Risk management is a name for a wide range

7. American Society of Civil Engineering (ASCE), *Guiding Principles for the Nation's Critical Infrastructure* (Reston, VA: ASCE, 2009), 15–16, https://www.asce.org/uploadedFiles/Issues_and_Advocacy/Our_Initiatives/Infrastructure/Content_Pieces/critical-infrastructure-guiding-principles-report.pdf.

8. Stanley Kaplan and B. John Garrick, "On the Quantitative Definition of Risk," *Risk Analysis* 1, no. 1 (1981): 13.

of methodologies for operating and maintaining a system in a manner that yields an acceptable level of risk to all relevant parties, including the system owners, operators, and regulators. See chapter 13 for an overview of risk assessment and management strategies. Effectively assessing and managing risk requires the identification of potential adverse events that could result in consequences for the system and its users.[9]

Unlike risk, *resilience* does not have a universal definition in the scholarly literature.[10] The definition of resilience varies by topic—economic resilience, critical infrastructure resilience, and social resilience, for example—and by object of analysis, such as community, infrastructure system, and infrastructure asset. Variations in the definition of resilience affect how decisionmakers measure and evaluate resilience.[11] In their thorough literature review on resilience definitions, L. Carlson et al. identified two major schools of thought, with the primary difference being whether actions taken prior to the occurrence of an adverse event would be considered as enhancing resilience.[12] Nearly all literature defines resilience, in part or in whole, as the ability to absorb a disturbance and adapt to or quickly recover from the changes caused by the disturbance to resume functionality.[13] Some literature further defines resilience as encompassing the ability to plan for, mitigate against, and withstand a hazard to reduce overall vulnerability and thus the likelihood of a threat or hazard causing a disruption.[14] By considering a subset of these pre-event actions that are taken under the assumption that an adverse event is going to occur, Carlson et al. defined resilience as "the ability of an entity—asset, organization, community, region—to anticipate, resist, absorb, respond to, adapt to, and recover from a disturbance."[15]

Risk and resilience are highly interrelated concepts; actions taken to reduce the risk of a system will likely increase the resilience of the system and vice

9. Kaplan and Garrick, "Definition of Risk," 13.

10. Stanley W. Gilbert, *Disaster Resilience: A Guide to the Literature* (Washington, DC: National Institute of Standards and Technology, 2010), 9–11, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication1117.pdf.

11. L. Carlson et al., *Resilience Theory and Applications* (Argonne, IL: Argonne National Laboratory, 2012), 17–20, https://publications.anl.gov/anlpubs/2012/02/72218.pdf.

12. Carlson et al., *Resilience Theory*, 15–17.

13. Joseph Fiksel, "Sustainability and Resilience: Toward a Systems Approach," *Sustainability: Science, Practice and Policy* 2, no. 2 (2006): 16, https://doi.org/10.1080/15487733.2006.11907980.

14. Department of Homeland Security (DHS), *Quadrennial Homeland Security Review Report: A Strategic Framework for a Secure Homeland* (Washington, DC: DHS, 2010), 59–64; and Adrian V. Gheorghe et al., *Critical Infrastructures, Key Resources, Key Assets* (Cham, CH: Springer International Publishing, 2018).

15. Carlson et al., *Resilience Theory*, 17.

versa. Assessing and informing risk reduction and resilience enhancement measures requires the identification of potential natural hazards and man-made threats as well as an assessment of system vulnerabilities, hazard mitigation actions, and response and recovery capabilities following a disruption event.[16]

These steps also necessarily implicate broader community and regional characteristics that the infrastructure serves. Infrastructure systems provide the commodities and services required by a population, either directly or indirectly. The seminal work on individual needs by A. H. Maslow conceptualized human needs as a pyramid with each of the lower levels serving as a foundation for higher-order development.[17] At the most foundational levels, Maslow's hierarchy describes the physiological and safety needs of individuals (such as water, food, sheltering, and security). These needs must be satisfied in order for individuals to find belonging, esteem, and self-actualization as matters of psychological development.[18]

A parallel construct that applies Maslow's hierarchy to societal development may help to define the questions of greatest importance for community and regional decision makers in formulating resilience strategies.[19] In this construct, various infrastructure systems (including water infrastructure, agricultural infrastructure, and emergency services) provide these basic societal needs.[20] In turn, these infrastructure systems require services and commodities provided by other infrastructure systems. For example, water treatment infrastructure, which processes raw water into safe, potable water, requires chemicals (such as chlorine and fluoride) provided by chemical infrastructure. As such, chemical infrastructure producing water treatment chemicals indirectly supports the needs of individuals. The quality of infrastructure systems and the goods and services they provide may have a strong influence on the quality of life and strength of a society.[21]

16. Frédéric Petit, "Resilience Assessment in Homeland Security," in *IRGC Resource Guide on Resilience: Domains of Resilience for Complex Interconnected Systems,* vol. 2, ed. Benjamin Trump, Marie-Valentine Florin, and Igor Linkov (Lausanne, CH: EPFL International Risk Governance Center, 2018): 119–20.

17. A. H. Maslow, "A Theory of Human Motivation," *Psychological Review* 50, no. 4 (1943): 370–96.

18. M. Joseph Sirgy, "A Quality-of-Life Theory Derived from Maslow's Developmental Perspective," *American Journal of Economics and Sociology* 45, no. 3 (1986): 329.

19. For additional information, see M. R. Hagerty, "Testing Maslow's Hierarchy of Needs: National Quality-of-Life across Time," *Social Indicators Research* 46, no. 3 (1999): 249–71.

20. For additional information, see Jochen Markard, "Infrastructure Sector Characteristics and Implications for Innovation and Sectoral Change," *Journal of Infrastructure Systems* 17, no. 3 (2011): 107–17.

21. For additional information, see Gheorghe et al., *Critical Infrastructures.*

Critical Infrastructure Interdependency Taxonomies and Concepts

The definitions of what constitutes a critical infrastructure are relatively consistent around the globe. Table 12-1 contains examples of critical infrastructure definitions used in the West.

Table 12-1. Definitions of critical infrastructure

Country	Definitions
Australia	Physical facilities, supply chains, information technologies and communication networks, which if destroyed, degraded, or rendered unavailable for an extended period, would significantly impact the social or economic well-being of the nation, or affect Australia's ability to conduct national defense and ensure national security.[22]
Canada	Processes, systems, facilities, technologies, networks, assets, and services essential to the health, safety, security or economic well-being of Canadians and the effective functioning of government.[23]
European Union	Asset or system that is essential for the maintenance of vital societal functions. The damage to a critical infrastructure, its destruction or disruption by natural disasters, terrorism, criminal activity, or malicious behavior, may have a significant negative impact for the security of the EU and the well-being of its citizens.[24]
United Kingdom	Those facilities, systems, sites, information, people, networks, and processes, necessary for a country to function and upon which daily life depends.[25]
United States	Infrastructure so vital that its incapacity or destruction would have a debilitating impact on defense and national security.[26]

Even if all these definitions recognized critical infrastructure as the main components supporting the well-being of the society, they would not reflect a consensus on which sectors comprise critical infrastructure. See chapter 1 for a discussion of the sectors generally identified as critical infrastructure.

22. Australian Government, *Critical Infrastructure Resilience Strategy* (Canberra: Commonwealth of Australia, 2010), 8.

23. "Critical Infrastructure," Public Safety Canada (website), n.d., accessed December 17, 2021, https://www.publicsafety.gc.ca/cnt/ntnl-scrt/crtcl-nfrstrctr/index-en.aspx.

24. "Critical Infrastructure," European Commission: Migration and Home Affairs (website), n.d., accessed December 17, 2021, https://ec.europa.eu/home-affairs/what-we-do/policies/crisis-and-terrorism/critical-infrastructure_en.

25. "Critical National Infrastructure," Centre for the Protection of National Infrastructure (CPNI) (website), April 20, 2021, https://www.cpni.gov.uk/critical-national-infrastructure-0.

26. President's Commission on Critical Infrastructure Protection, *Critical Foundations: Protecting America's Infrastructures* (Washington, DC: White House, 1997), 3, https://www.hsdl.org/?abstract&did=986.

Table 12-2 provides examples of the various ways in which Western countries organize and classify sectors as critical infrastructure.

Table 12-2. Critical infrastructure sector taxonomy

Country	Sectors
Australia	7 sectors: energy, water services, communications, transport, food chain, health, banking and finance[27]
Canada	10 sectors: health, food, finance, water, information and communication technology, safety, energy and utilities, manufacturing, government, transportation[28]
European Union	2 sectors: energy (i.e., electricity, oil, and gas), transport (road, rail, air, inland waterways, and ocean and short sea shipping and ports)[29]
United Kingdom	13 sectors: chemical, civil nuclear, communications, defense, emergency services, energy, finance, food, government, health, space, transport, water[30]
United States	16 sectors: chemical, commercial facilities, communications, critical manufacturing, dams, defense industrial base, emergency services, energy, financial services, food and agriculture, government facilities, health care and public health, information technology, transportation systems, water and wastewater systems, and nuclear reactors, materials, and waste[31]

All these taxonomies include at least the lifeline sectors—energy, water, communications, and transportation—that support the physiological needs defined by Maslow. The diversity of characterizations of critical infrastructure sectors is symptomatic of the complexity of these systems and of their strong interconnections.

Critical infrastructure assets constitute a system of systems in which resources—namely goods or services—supplied by one network constitute the raw materials supporting the operations of other networks. Figure 12-1 provides a high-level example of interdependencies among lifeline infrastructure systems.[32]

27. Australian Government, *Critical Infrastructure Resilience*, 10.

28. "Critical Infrastructure," Public Safety Canada.

29. "Critical Infrastructure," European Commission.

30. "Critical National Infrastructure," CPNI.

31. White House, *PPD-21.*

32. Duane Verner, Agnia Grigas, and Frédéric Petit, *Assessing Energy Dependency in the Age of Hybrid Threats* (Helsinki: European Centre of Excellence for Countering Hybrid Threats, 2019), 5, https://www.hybridcoe.fi/wp-content/uploads/2019/02/Assessing_Energy_Dependency_in_the_Age_of_Hybrid_Threats-HybridCoE.pdf.

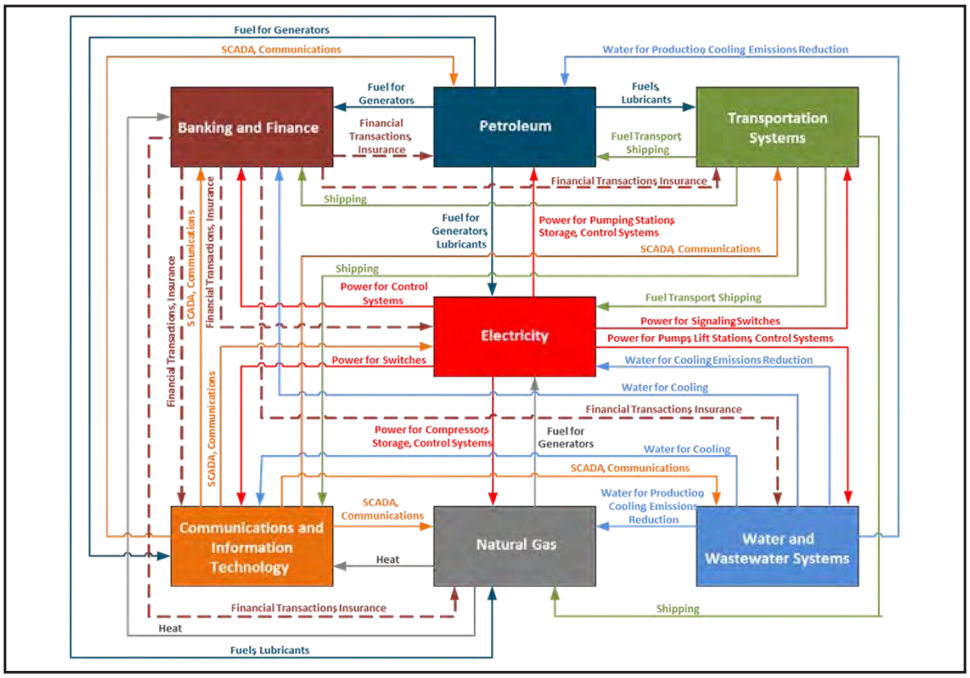


Figure 12-1. Interdependencies among lifeline systems
(Diagram by Argonne National Laboratory)

As illustrated in figure 12-1, interdependencies are characterized by the exchange of resources constituting both the inputs and outputs of different infrastructure assets. Therefore, understanding interdependencies existing within and among critical infrastructure systems involves characterizing the relative importance of upstream, internal, and downstream dependency categories. Figure 12-2 illustrates the interaction between a critical infrastructure and its environment.[33]

33. Frédéric Petit et al., *Analysis of Critical Infrastructure Dependencies and Interdependencies* (Argonne, IL: Argonne National Laboratory, 2015), 9, https://doi.org/10.2172/1184636.

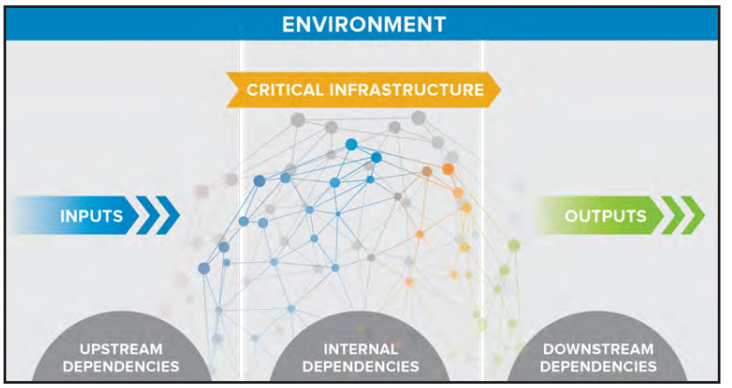


Figure 12-2. Critical infrastructure and its environment
(Diagram by Argonne National Laboratory)

The operation of a critical infrastructure system requires commodities and services provided by other infrastructure systems. *Dependency* can be defined as a unidirectional link between two infrastructure assets in which the service or commodity provided by the first asset to the second asset is necessary for operations of the latter.[34] An *interdependency* can be characterized as a combination of two dependencies; it is a bidirectional link between two assets.[35] In practice, understanding an interdependency requires understanding the two dependencies between the two assets comprising the interdependency.[36] Figure 12-3 illustrates the concepts of dependency and interdependency.[37]

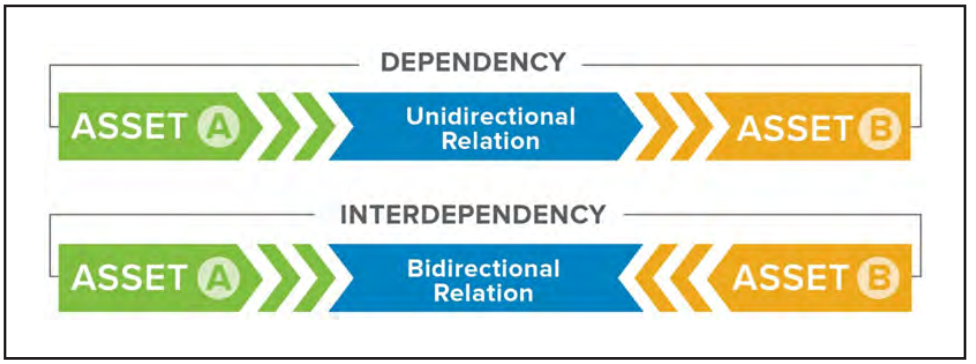


Figure 12-3. Dependency and interdependency between assets
(Diagram by Argonne National Laboratory)

34. S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies," *IEEE Control Systems Magazine* 21, no. 6 (2001): 13–14, https://doi.org/10.1109/37.969131.

35. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 14.

36. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 14.

37. Petit et al., *Analysis of Critical Infrastructure*, 8.

The types of dependencies and interdependencies may be of a different nature and have specific characteristics. Several authors have developed taxonomies for infrastructure interdependencies. The classification developed by S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, however, is the most comprehensive and the most widely used for critical infrastructure analysis. This taxonomy defines the four classes of dependencies listed in table 13-3.[38]

Table 13-3. Dependency classes

Classes	Definition
Physical	Operations depend on material output(s) of other infrastructure through a functional and structural linkage between the inputs and outputs of two assets.
Cyber	Operations depend on information and data transmitted through the information infrastructure via electronic or informational links.
Geographic	Operations depend on the local environment, where an event can trigger changes in the state of operations in multiple infrastructure.
Logical	Operations depend on the state of other infrastructure via connections other than physical, cyber, or geographical. Logical dependency is attributable to human decisions and actions and is not the result of physical or cyber processes.

Rinaldi, Peerenboom, and Kelly argue that these four classes are not the only characteristics to consider when assessing critical infrastructure dependencies. Table 13-4 lists several other dimensions that the assessment process must take into account.[39]

Table 13-4. Dependency dimensions

Dimensions	Definition
Operating environment	Characterize how external factors can influence infrastructure operations and connections
Coupling and response behavior	Characterize how an infrastructure would react to infrastructure connection disruptions or changes
Type of failure	Characterize the propagation of consequences resulting from a disruption
Infrastructure characteristics	Characterize the infrastructure organization and operation
State of operation	Characterize the critical infrastructure state of operations

38. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 14–16.

39. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 16–23.

Enhancing the resilience of critical infrastructure requires a better understanding of the interdependencies among critical infrastructure assets and the influence of these interdependencies on the operations of different infrastructure systems. This level of understanding requires consideration of the categories, classes, and dimensions that characterize infrastructure dependencies and interdependencies. Integrating the characterization of infrastructure interdependencies into an infrastructure's risk management process enables an understanding of how interdependencies influence all components of risk: threats, vulnerabilities, resilience, and consequences. Interdependencies can increase the intensity of man-made threats and natural hazards. They can also expand the set of facility vulnerabilities, which are those vulnerabilities that a facility may be unable to address effectively because interdependency elements are often outside a facility's control. Interdependencies also influence mitigation and response capabilities. Finally, interdependencies are the main factor influencing propagation of consequences—cascading and escalating failures—across infrastructure systems and regions.

Cascading failures (or *domino effects*) represent the succession of disruptions within an infrastructure system and across infrastructure systems.[40] For example, the failure of a transformer in an electric power distribution substation may make it impossible to operate the substation, and ultimately lead to a power outage in the substation service area. The loss of electricity will potentially impact other infrastructure assets located in this area and therefore affect the operations of their respective infrastructure systems. *Escalating failure* (or *snowball effect*) represents an increase of severity or time to respond to an existing infrastructure failure.[41] For example, considering the dysfunction of the electric power distribution substation, if the loss of power affects the transportation system—such as dysfunctions of traffic control centers and traffic lights—these problems can result in a delay for the repair crew to access the substation to replace the failed transformer.

Even though research addressing critical infrastructure interdependencies started in the United States more than 20 years ago with the President's Commission on Critical Infrastructure Protection, modeling, simulation, and visualization tools still perform at an intermediate analysis level.[42]

40. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 22.

41. Rinaldi, Peerenboom, and Kelly, "Critical Infrastructure Interdependencies," 22.

42. See Frédéric Petit et al., "Critical Infrastructure Protection and Resilience—Integrating Interdependencies," in *Security by Design: Innovative Perspectives on Complex Problems*, ed. Anthony J. Masys (Cham, Switzerland: Springer International Publishing, 2018), 193–219.

Data collection and analyses generally operate in isolation to address first-order physical and cyber dependencies.[43]

Critical Infrastructure Modeling

Given the importance of critical infrastructure systems in supporting the needs and operation of communities, analyses of critical infrastructure systems are vital to ensuring the resilience and health of communities.[44] Infrastructure interdependency assessments, however, can be analytically complicated, time consuming, and costly, which in turn can limit stakeholders' abilities to understand and use this information to make risk-informed decisions.

Numerous infrastructure analysis methodologies have been developed to provide infrastructure owners and operators with various information. These methodologies—including infrastructure vulnerability, risk, and resilience assessments, infrastructure expansion studies, and infrastructure interdependency analyses—help inform infrastructure investment planning, business continuity planning, and operational decision making.

Generally, risk assessment and system engineering methodologies are also useful for analyzing critical infrastructure interdependencies. Examples of common methodologies for assessing infrastructure systems include:

- Network modeling and graph theory, including Voronoi or Thiessen polygons, Huff modeling, and cellular automata[45]
- System of systems modeling, such as multilayer infrastructure network[46]
- Simulation-based modeling, such as agent-based modeling[47]

43. Petit et al., *Analysis of Critical Infrastructure*, 24–25.

44. Carlson et al., *Resilience Theory and Applications*, 31–38.

45. For additional information, see Irene Eusgeld et al., "The Role of Network Theory and Object-oriented Modeling within a Framework for the Vulnerability Analysis of Critical Infrastructures," *Reliability Engineering & System Safety* 94, no. 5 (2009): 954–63, https://doi.org/10.1016/j.ress.2008.10.011.

46. For additional information, see Pengcheng Zhang and Srinivas Peeta, "A Generalized Modeling Framework to Analyze Interdependencies among Infrastructure Systems," *Transportation Research Part B: Methodological* 45, no. 3 (2011): 553–79, https://doi.org/10.1016/j.trb.2010.10.001.

47. For additional information, see Michael J. North, "Multi-agent Social and Organizational Modeling of Electric Power and Natural Gas Markets," *Computational & Mathematical Organization Theory* 7 (2001): 331–37, https://doi.org/10.1023/A:1013406317362.

- Economic modeling, including input-output modeling[48]
- Multi-criteria decision analysis[49]
- Game theory[50]

Infrastructure interdependency modeling methodologies can be generally categorized as mathematics, geospatial, system dynamics, economics, or physics-based. Developed for specific types of critical infrastructure or for specific conditions, each of the modeling and simulation approaches presents advantages and limitations. These models, which can be deterministic or probabilistic in nature, may evaluate the performance of an infrastructure system at a single point in time (static) or over a given period (dynamic). Each infrastructure modeling methodology has advantages and disadvantages, including data requirements, assumptions, level of granularity, model complexity, and model development and execution times. No single model type is best for all types of infrastructure assessments. Rather, infrastructure analysts should weigh the advantages and disadvantages of different modeling methodologies against the needs of the stakeholders and limitations of the analysis when selecting an infrastructure modeling methodology.

Analysis of infrastructure systems conventionally occurs in a siloed manner, which means that directly modeling the assets within the infrastructure system is the basis for the evaluation of the system's operation. If the traditional siloed approach considers dependencies and interdependencies, it is only in an indirect manner, such as via input models or explanatory variables informed by expert opinion, elicitation, or historical data. Assessing infrastructure in a siloed nature significantly inhibits a broader understanding of the dynamics between different infrastructure systems, including cascading and escalating failures and impacts to supply chains. By explicitly modeling dependencies using a system of systems approach, analysts can gain a better understanding of the interconnected operations of infrastructure systems, which may enable the identification of single points of failure that conventional siloed analyses may not identify. Recent works characterize interdependencies across infrastructure systems, but they generally focus on lifeline networks and require the support of organizations with research and development

48. For additional information, see Y. Haimes and P. Jiang, "Leontief-Based Model of Risk in Complex Interconnected Infrastructures," *Journal of Infrastructure Systems* 7, no. 1 (2001): 1–12.

49. For additional information, see Darwin Marcelo et al., *Prioritizing Infrastructure Investment: A Framework for Government Decision Making* (Washington, DC: World Bank Group, 2016), https://openknowledge.worldbank.org/handle/10986/24511.

50. For additional information, see Pengcheng Zhang, Srinivas Peeta, and Terry Friesz, "Dynamic Game Theoretic Model of Multi-layer Infrastructure Networks," *Networks and Spatial Economics* 5, no. 2 (2005): 147–78, https://doi.org/10.1007/s11067-005-2627-0.

capabilities.[51] Furthermore, most of the existing methodologies focus primarily on physical interdependencies and, to some extent, on cyber dependencies; few works incorporate geographic or logical dependencies.[52]

Approaches to infrastructure modeling typically fall into one of two categories: top-down or bottom-up. Top-down modeling approaches focus on the operation of the entire infrastructure system, and then estimate the performance and significance of infrastructure subsystems and assets based on the overall system dynamics. Bottom-up modeling approaches focus on the attributes and operation of the individual assets comprising an infrastructure system, and then estimate the overall operation of the system based on the performance of the individual assets. Top-down modeling may provide greater accuracy of the overall system operation with less detail on the functioning of individual assets within the system. Conversely, bottom-up approaches provide a high level of detail on the functioning of individual assets but provide less information on overall system operation. Combining top-down and bottom-up approaches can yield high levels of detail on the operation of both the overall system and the assets comprising the system.

A system of systems approach combining top-down and bottom-up assessments within and across critical infrastructure systems helps to establish the appropriate scope of interdependency analyses and identify specific security and resilience management strategies.[53] Figure 12-4 illustrates how combining different analysis approaches can better characterize critical infrastructure interdependencies.[54]

51. For additional information, see Constantinos Heracleous et al., "Hybrid Systems Modeling for Critical Infrastructures Interdependency Analysis," *Reliability Engineering & System Safety* 165 (2016): 89–101.

52. For additional information, see Petit et al., *Analysis of Critical Infrastructure.*

53. Igor Linkov et al., "Risk-based Standards: Integrating Top-down and Bottom-up Approaches," *Environment Systems and Decisions* 34 (2014): 134–37, https://doi.org/10.1007/s10669-014-9488-3.

54. Frédéric Petit, Duane Verner, and Leslie-Anne Levy, *Regional Resiliency Assessment Program Dependency Analysis Framework* (Argonne, IL: Argonne National Laboratory, 2017), 9, https://doi.org/10.2172/1475551.

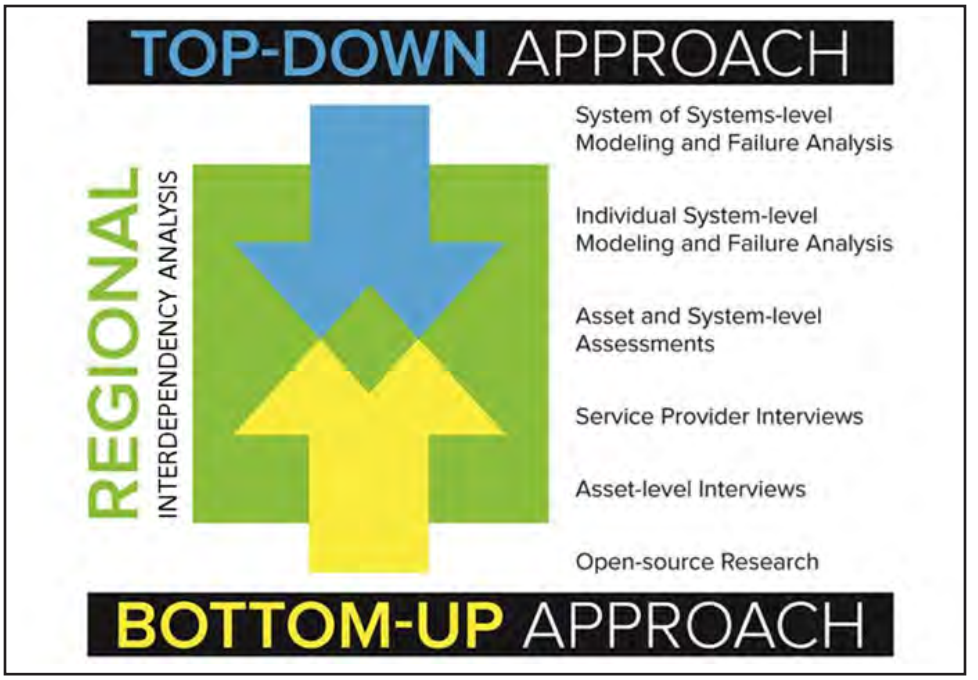


Figure 12-4. Resilience interdependency analysis overview
(Diagram by Argonne National Laboratory)

Developing a systemic approach to move toward an advanced infrastructure analysis capability presents several challenges.[55] These challenges include:

- Lack of understanding of infrastructure system operations and of the interactions existing between infrastructure systems.
- Lack of understanding of how infrastructure interdependencies affect the resilience of a region.
- Difficulty in combining simulation models originally developed in silos and specific to certain critical infrastructure systems.
- Difficulty in obtaining the data for running the simulation models. Some data are proprietary and some are unknown due to the lack of knowledge.

Promoting a collaborative approach based on trust and information sharing can overcome most of these challenges and the intrinsic complexities of critical infrastructure interdependencies.

55. Clifford and Macal, *Advancing Infrastructure Dependency*, 2–7.

Critical Infrastructure Interdependency Analysis Framework

The significance of critical infrastructure systems in supporting the needs of the public and broader communities and regions underlies the need for a standardized methodology for systematic assessments of infrastructure systems. Given the diversity of infrastructure systems and assets, and the breadth of infrastructure analysis methodologies, figure 12-5 proposes a generalized critical infrastructure interdependency analysis framework.

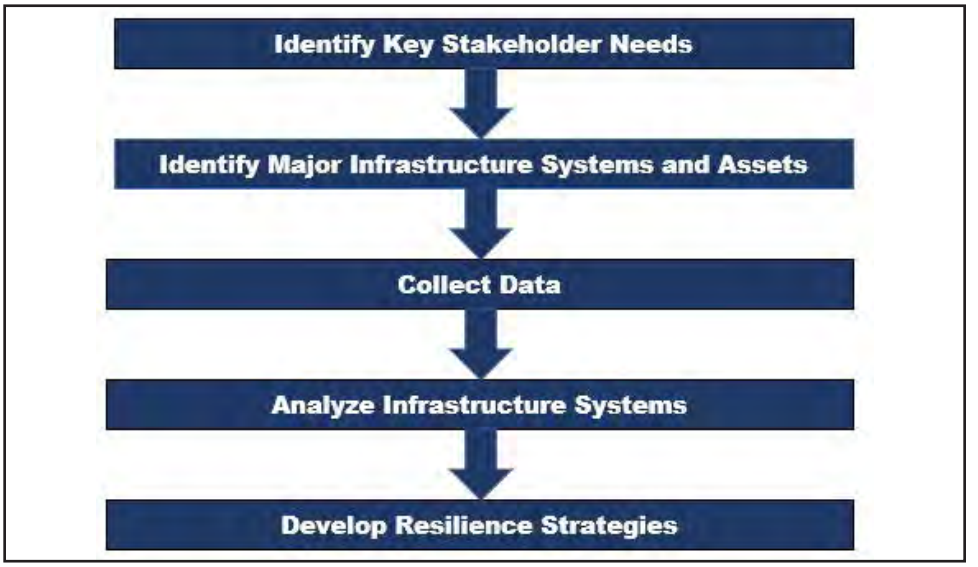


Figure 12-5. Interdependency analysis framework

Identification of Key Stakeholders' Needs

A starting point is to determine the key stakeholders, such as infrastructure system owners and operators, government agencies, nongovernmental organizations, and the public. Then, the first step in the critical infrastructure interdependency analysis framework is to identify the operational and organizational needs of these key stakeholders. Key stakeholder needs include both upstream dependencies, such as utilities providing critical resources, and downstream dependencies, including customers and other infrastructure systems that require the commodities or services provided by key stakeholders. Although dependency analyses generally focus on the impact of the loss of upstream dependencies, it is equally important to consider the impact of the loss of an infrastructure system on downstream users and the community and societal environment in which the system operates. Following a disruption event, the shifting needs of downstream users and the operating environment

can—and, perhaps, should—significantly influence the operation of key stakeholders' infrastructure.[56]

One of the primary goals of the framework is to provide key stakeholders the information needed to inform decision making for improving infrastructure resilience. As such, it is important to identify and delineate the decision-making process of the key stakeholders in order to identify the types and granularity of data requirements for informing stakeholder decision making. Part of this step involves initial stakeholder outreach and engagement in order to open lines of communication and garner support from key stakeholders and the critical infrastructure owners and operators managing infrastructure that are upstream dependencies of primary stakeholders' infrastructure.

Identification of Major Assets and Systems

Based on specific stakeholders' needs, their desired level of information granularity, and the particular social systems and communities in which their infrastructure systems operate, the second step is to identify the relevant infrastructure systems and assets of interest. A key part of this step is the prioritization of the assets within the primary stakeholders' infrastructure systems and within the infrastructure systems that are upstream dependencies of these systems. Prioritizing infrastructure assets and systems should reflect the expected impacts to the operation of the stakeholders' infrastructure and to the downstream users and customers following the disruption or damage of the assets or systems. Part of the goal of this step is to guide the data collection process.

Data Collection

The third step consists of collecting data on the prioritized infrastructure assets and systems in order to characterize the upstream and downstream dependencies of key stakeholders. Publicly available, open source datasets—published by governmental, nonprofit, educational, trade, research, and other nongovernmental organizations—are a good starting point for data collection. Infrastructure data collected by private organizations may be available for purchase in proprietary datasets. Restrictions on the use of these datasets, however, may prohibit the sharing of pertinent information and analysis results with key stakeholders, thus limiting the utility of the data. Before purchasing proprietary datasets, it is important to evaluate carefully the limitations on the use of proprietary data. Beyond combing public and proprietary datasets

56. Roberto Guidotti, Paolo Gardoni, and Nathanael Rosenheim, "Integration of Physical Infrastructure and Social Systems in Communities' Reliability and Resilience Analysis," *Reliability Engineering & System Safety* 185 (May 2019): 476–78, https://doi.org/10.1016/j.ress.2019.01.008.

for relevant infrastructure data, the collection process also involves requesting data directly from key stakeholders and the utilities that provide them essential commodities. Additionally, surveys, site visits, facilitated discussions, and structured interviews with key stakeholders are all viable means to collect data. Data of interest includes information on infrastructure vulnerabilities, dependencies, and operations, as well as information on the political and socioeconomic environment in which these infrastructure assets exist and operate.

Infrastructure Analysis

Once all pertinent data is collected, the next step is to evaluate the performance of infrastructure systems. The specific needs of the key stakeholders identified in the first step of the framework informs which type of infrastructure analyses to conduct in this fourth step. System-level (top-down) analysis methodologies can provide information on the operation of infrastructure systems and can assess the cascading failures in an infrastructure system following the disruption of assets within the system. Asset-level (bottom-up) approaches can assess the vulnerabilities and resilience of specific infrastructure assets. Combining system- and asset-level approaches enables assessment of the cascading and escalating failures across infrastructure systems following the disruption of infrastructure assets, and facilitates the identification of critical infrastructure assets to prioritize for resilience enhancements and investments.

Definition of Resilience Strategies

The fifth step of the framework is the development of strategies to address resilience deficiencies identified throughout all other steps of the framework. Resilience measures can address the vulnerabilities and dependencies of critical assets owned by key stakeholders and utilities. This final step seeks to propose resilience strategies applying at both asset and regional levels. When key stakeholders have a major vulnerability due to upstream dependencies, resilience strategies often involve evaluation of potential alternative or back-up capabilities as well as increased communication with the relevant utilities and infrastructure owners and operators. One of the primary objectives of this step is to disseminate important findings and proposed resilience strategies via out briefs, public forums, facilitated discussions, workshops, presentations, reports, and other communication methods.

Operationalization of Critical Infrastructure Interdependencies

In order to be effective, risk reduction efforts should target the broader enhancement of community or regional resilience, which requires the consideration of interdependencies within and across infrastructure systems. The proposed framework leverages classical risk management approaches and may appear simplistic for considering all elements—that is, categories, classes, and dimensions—characterizing critical infrastructure interdependencies. This simplicity is also a strength of the framework. Analysts are familiar with this specific assessment structure, and it provides sufficient flexibility to adapt the level of assessments to the users' needs and analysts' capabilities.

To be truly effective, the analysis framework requires:

- A working environment based on trust.
- Processes to deal with sensitive information.
- Community or regional coordination.

Developing trust among key stakeholders and analysts is necessary to promote a collaborative and multidisciplinary working environment. Diverse participation also enables the analysis to incorporate social, economic, and technical considerations. The process is fully effective when it considers all key stakeholders involved in critical infrastructure management and emergency management, including the public. Beyond developing trust, it is also important to adopt mechanisms to operationalize standards and policies that promote collaborative approaches and partnerships between critical infrastructure owners and operators, and government representatives.

When the environment of trust is established, communication mechanisms must be implemented to maintain a balance between protecting sensitive information—from a business sensitivity and/or national security perspective—and providing key stakeholders, emergency managers, and government agencies necessary information to support resilience management strategies. See chapter 11 for a discussion of information and intelligence sharing principles involving key stakeholders. Understanding regional capabilities is paramount for coordinating resilience strategies. Malevolent actors, however, can exploit infrastructure interdependencies information to create security vulnerabilities. By identifying system weaknesses and admitting

their system can fail, infrastructure owners and operators could also generate a loss of public confidence, potentially leading to serious economic ramifications.

Community or regional coordination is important for defining what constitutes an acceptable level of consequences, not only for the critical infrastructure assets themselves, but also for the entire area served by the critical infrastructure systems. Understanding critical infrastructure interdependencies and defining tolerable levels of disruptions is important to prioritize protection, mitigation, response, and recovery efforts. Securing critical infrastructure, especially in complex urban areas often targeted by terrorists, should focus on identifying and prioritizing potential failure points that would have the most severe consequences. Community or regional coordination, combined with criticality analysis, will help infrastructure system owners and operators and government agencies identify priority assets for in-depth security and resilience assessments, and inform resilience investment decisions.

Conclusion

Infrastructure interdependencies are important elements for NATO member states and partner nations to consider for managing critical infrastructure systems, enhancing their security and resilience, and maintaining societal functions. Several elements influence infrastructure interdependencies, from the different classes of dependencies to the characteristics of their socioeconomic environment. Even though modeling critical infrastructure interdependencies may appear overwhelming, it is possible to assess these complex and dynamic relationships by combining top-down and bottom-up assessment techniques in an adaptive and flexible analysis framework. Building on a collaborative process promoting information sharing and prioritizing critical infrastructure assets, the assessment framework can facilitate a better understanding of infrastructure operations and help stakeholders anticipate and prepare for potential cascading and escalating failures.

The objective of assessing critical infrastructure interdependencies is to go beyond traditional, isolated risk assessment and management approaches. The framework proposed in this chapter may help analysts to move toward the development of coordinated resilience strategies that integrate key stakeholder needs, diverse infrastructure data, and the combination of analytical techniques—including inter-Alliance modeling and stress testing—that will ultimately guide more effective decision making.

In light of the risks posed by the complex natural, man-made, and socio-technical hazards NATO faces, infrastructure interdependency analysis will equip member states and partner nations with a transdisciplinary and multidimensional understanding of how critical infrastructure operates. Leveraging this deeper understanding of how these infrastructure assets and systems work in concert—both under normal and stressed conditions caused by terrorism—will be essential in order to build and maintain the long-term viability of NATO into the twenty-first century.

— 13 —

Security Risk Assessment and Management

Geoffrey French

How can the North Atlantic Treaty Organization best manage and assess security risk in a constantly changing environment? In some ways, security risk management is an intuitive and natural set of judgments that people make every day. Using information on the weather or current events to inform decisions about whether to carry an umbrella, bring certain valuables, or cancel an activity entirely, for example, is a common way to consider the things that may happen and take precautions against them. Complex, long-term, or high consequence decisions cannot rely on intuitive or ad hoc processes, however, with the hope of consistently positive results. Organizations and communities need formal processes to determine what the risks are, prioritize them, and address them. Continuous improvement in these processes and their results requires documentation and cyclic review. These seemingly simple concepts can become complex as they address complicated risks, such as natural disasters and terrorism, but they underpin the discipline of security risk management nonetheless. This chapter explores these concepts and how they translate to governmental programs for securing critical infrastructure against a range of threats, highlighting the role that a national-level risk program should play to coordinate the constellation of public- and private-sector organizations involved in critical infrastructure operations, and emphasizing the necessary characteristics of high-quality national risk programs.

Defining Security Risk Management

In the mid-1990s, the United States established a Joint Security Commission to review the decision-making processes that underpin governmental physical and information security investments. The commission noted the Cold War mentality had led to security investments based on unrealistic assumptions of the threat and an emphasis on risk elimination. The result of this mentality was that security measures were ad hoc, difficult to justify, and very expensive. The commission recommended an approach to balance the risk of loss or damage against the costs of countermeasures: a "rational, cost-effective, and enduring framework using risk management as the underlying basis for security decision-making."[1] In many ways, the vision established by the commission remains as an ideal, not yet achieved or perfected. The commission's work remains relevant decades later because the challenges it identified persist.

The US Government Accountability Office (GAO)—which supports Congressional oversight of executive branch agencies and departments, among other responsibilities—considers security risk management to be a fundamental part of decision making. The GAO has said that implementing "principles of risk management can help policymakers reach informed decisions regarding the best ways to prioritize investments in security programs so that these investments target the areas of greatest need."[2] In this sense, security risk management is an essential part of planning and communication. It allows leadership to see the competing security needs and prioritize the investments in policies, people, equipment, or systems. Communicating these priorities in strategic planning and organizational activities helps set the culture for an agency, location, or community.

Part of that communication is the adoption and consistent use of terms and their definitions. There are many ways to define *risk*. The International Organization for Standardization (ISO) defines risk as "the effect of uncertainty on objectives."[3] Although this ISO definition has the benefit of wide applicability to many fields, it may be too general to allow simple adoption in a security context. NATO's Science and Technology Organization defines risk for qualitative risk assessment

1. Joint Security Commission, *Redefining Security: A Report to the Secretary of Defense and the Director of Central Intelligence* (Washington, DC: Government Printing Office, 1994), 4.

2. Norman J. Rabkin, *Risk Management: Strengthening the Use of Risk Management Principles in Homeland Security* (Washington, DC: Government Accountability Office, 2008), 1.

3. International Organization for Standardization, *ISO 31000: Risk Management Guidelines* (Geneva: International Organization for Standardization, 2018), 1.

as "the potential of loss to an organization or entity."[4] Similarly, NATO's Allied Joint Publication on force protection defines risk as "a function of the value of the asset and . . . compared to the potential impact of the exploitation of vulnerabilities by threats and hazards."[5] The important elements of all of these definitions is that security risk considers the relative likelihood—including both threat and vulnerability—and consequences of unwanted events.

These terms are especially important in the realm of critical infrastructure security and resilience (CISR) because CISR involves a consortium of organizations: military and civilian and public and private, each with its own culture and vocabulary. The diversity of types of infrastructure to protect—and the types of threats that demand consideration—presents a dizzyingly complex situation for governmental and industry leaders. Clear communication, therefore, requires clear terminology. Language is not enough. The GAO states that *risk management* is a "strategic process for helping policymakers make decisions about assessing risk, allocating finite resources, and taking actions under conditions of uncertainty."[6] Leaders need a deliberate analytic approach to help them understand the short- and long-term risks so they can make informed decisions about resource allocation. Security risk analysis can serve as that approach for an organization and its stakeholders to the extent that the organization uses sound methods, implements repeatable processes, and documents its data and assumptions.

Risk Management Frameworks

The promise of risk management is that with sufficient uniformity and consistency, leaders can make better decisions through the ability to aggregate risks at different levels. That is, if tactical risk assessments are compatible, then a leader can begin to characterize the risk at the operational or strategic level. The inherent difficulty, however, is that threats, attack methods, and infrastructure assets and systems can differ widely. An overly specific set of instructions for how to measure risk in one area can be meaningless in another. The approach to assessing the risk of an explosive attack on a building by terrorists is very different in terms of data and analysis from the risk of a cyberattack on intellectual property by a foreign intelligence service. For these reasons, the CISR community has long relied on security

4. North Atlantic Treaty Organization (NATO) Science and Technology Organization, *Mission Assurance for Autonomous Unmanned Systems* (Brussels: NATO, 2018), 7-5.

5. NATO Standardization Office, *Allied Joint Doctrine for Force Protection*, Allied Joint Publication 3.14 (Brussels: NATO, 2015), B-3.

6. Rabkin, *Risk Management*, 1.

risk frameworks, rather than detailed guidance. Properly constructed, a framework presents the irreducible minimum of parts or characteristics of a risk management process that ensure due diligence rather than specifies a model or approach.

Table 13-1 presents the specific steps described in a set of selected risk frameworks—from the ISO, NATO, GAO and the US National Infrastructure Protection Plan (NIPP)—that have been designed or adapted for security risk management. There are many similarities across the frameworks, and individual frameworks may combine one or more steps that other models separate. Their commonalities show the key areas of agreement (as these frameworks have developed over the years) and how the ISO model makes some steps explicit that other models imply. For example, communicating the results of the risk assessment and the decision processes are crucial to a coherent program, but the ISO framework identifies it as a specific phase or action to emphasize its importance. Seeing the steps side-by-side clarifies the minimum phases, including those that some models combine or imply.

Table 13-1. Comparison of phases outlined in select risk frameworks

Risk Frameworks				
Phases	**GAO**	**NIPP**	**NATO**	**ISO**
Purpose and context	1. Strategic goals, objectives, and constraints	1. Set security goals		1. Establish context
Screening and scope		2. Identify assets, systems, networks, and functions	1. Identify hazards and threats	2. Risk identification
Risk assessment	2. Risk assessment	3. Assess risks	2. Assess hazards and threats	3. Risk analysis
Risk management analysis	3. Alternatives evaluation	4. Prioritize	3. Develop controls	4. Risk evaluation
Management decision	4. Management selection		4. Implement controls	5. Risk treatment
Implementation	5. Implement and monitor	5. Implement protective programs		
Monitoring		6. Measure effectiveness	5. Supervise and evaluate	6. Monitoring and review
Communication				7. Communication

A crucial aspect to a functional risk framework is its cyclic nature. First, it builds organizational maturity. As organizations become increasingly familiar with the analysis and management cycle, they become more attuned to the need for information and data and shared understanding of the their own goals. As an organization assesses a risk multiple times over the years, it should obtain or generate the data required for a clearer understanding of that risk. In other words, implementing a cyclic risk framework emphasizes the need for evidence-based decision making and prioritizes institutional learning and accumulation of information for the underlying evidence base.

Second, it encourages a culture of continuous improvement. The periodic review of risk management decisions communicates that management is a process that must be maintained. Evaluating past investments helps clarify their purpose by linking them to metrics or criteria that determine success. This evaluation highlights the need for equipment maintenance, personnel training, and the periodic refreshing of procedures.

Third, a cycle of analysis does not guarantee an organization will detect emerging threats. It creates the conditions for a regular scan and evaluation of emerging threats (such as the hybrid threats discussed in chapter 4). The intent is not to generate an unceasing list of risks that require investment, but rather an organizational understanding of the highest priority current, future, and potential risks.

National Risk Programs

Every nation faces a range of security threats that must be addressed. Many—if not all—of these threats require specialized knowledge, skills, and data that drive organizational design to focus on single risks or a narrow range of risks. It makes sense, for example, to establish centers that study hurricanes, wildfires, earthquakes, or other natural hazards, as well as law enforcement, intelligence, or academic centers focused on organized crime or terrorism. Industry, too, has specialization by sector and subsector, and, in some cases, specific technologies used in processes (such as industrial control systems or supervisory control and data acquisition systems). The different types of organizations responsible for CISR represent nodes of expertise or knowledge sets, and as they mature, they develop links among them. See chapter 10 for a discussion of the various stakeholders and organizations involved in CISR. They are unlikely to create a coherent network

of nodes and links without deliberate planning and coordination, which should be one of the primary roles of a risk program at the national level.

A national risk program not only entails producing risk analysis, but also guiding and encouraging the generation and aggregation of data useful for risk analysis, as well as providing the frameworks necessary to compare and combine analysis where appropriate.[7] In this way, a national risk program can serve as a central hub to assist public and private partners with authoritative and reliable sources for either data or analysis. In some cases, the government can serve as a conduit or convening authority that allows and encourages information sharing among private sector entities. See chapter 11 for an overview of the foundational concepts that should underpin CISR information-sharing programs.

By building trusted partnerships and a network of credible data, the national risk program supports evidence-based policy making. Risk frameworks enable information from one field to be applied in another. A clear example of this trust is enabling the translation of intelligence, law enforcement, and open-source analysis to structured threat analysis, so that public and private partners understand the relative threat levels and how to incorporate them into risk analysis appropriately. This partnership requires an understanding and respect for the specialists and a willingness to serve as an ambassador, guiding specialists so their analysis is useful to the other stakeholders. Threat analysis that is too vague (or that does not further the understanding of a group's intent, capability, or targets) has very limited value to the organizations that require insight into the threats to complete their risk assessments.

Beyond making connections among public and private sector partners, a national risk program must evaluate the aggregate value of those connections and determine where there are gaps in metrics, data, analytic capability, or assessments and then coordinate action to address the gaps. Any segmentation of infrastructure, whether bottom-up or top-down, is likely to have gaps or contain areas of overlap. These areas require a risk governance process to oversee the identification of the gaps and overlaps, their prioritization, and their amelioration. In some cases, these gaps and overlaps may have a geographic component as well, where an infrastructure function is spread among jurisdictions and no single part is a local government priority even though the overall function is critical, or where

7. Edward J. Jopeck and Kerry L. Thomas, "Security Risk Management: Implementing a National Framework for Success in the Post 9/11 World" in *Critical Infrastructure Protection: Elements of Risk* (Arlington, VA: George Mason University School of Law, 2007), 1.

one jurisdiction benefits from an infrastructure function but a separate jurisdiction incurs some cost without receiving the benefit.

Finally, a national risk program should also play a strategic role in managing risk. Managing risks at the national level begins with fostering the ability of local government and infrastructure sectors to assess risk, building on the steps outlined above. Although the foundation is information sharing, a national risk program has a responsibility to provide or guide the development of tools that assist with coherent and consistent risk assessments. The benefit is not only to the local governments and sector-specific entities, but also to the nation. As local governments build comparable risk assessments, the national risk program can obtain a more thorough understanding of the types of risks the nation faces, differences in perception, and variances in risk management. Sharing insights and best practices propagates effective responses and minimizes less effective or negligent approaches. This bottom-up approach can also inform national threat analysis; local government and law enforcement are likely to encounter members of criminal and terrorist organizations, see the effects of their operations, or receive suspicious activity reports. Local insights can be highly informative to building the national-level perspective.

Although national risk management begins with a risk governance function, ensuring that risks are properly identified, prioritized, and managed, it does not end there. A national-level entity is best positioned to see the potential for collective action. In some cases, these actions may be major capital investments, but often they may be policy mechanisms that incentivize behaviors that maximize community resilience and minimize short-sighted actions or investments that benefit a small number of organizations and reduce the resilience of the infrastructure sector or surrounding community. See table 13-2 for examples of governmental risk controls.

Table 13-2. Government risk controls in the critical infrastructure environment

Organizations seeking to control risks may invest in facilities, equipment, or people, among other options. Government programs, for example, can build governmental capabilities in threat detection, specialized response, or communications. CISR, however, is a collective effort that involves individual organizations making decisions that affect others. The related governmental agencies can work together to influence the collective effort but are rarely in a position to direct investments or action. Governmental influence often comes through policy, guidance, standards, and regulations. Two primary challenges in this area are: (1) providing effective incentives, and (2) holding individual organizations accountable when they have not addressed security risks in a responsible manner. Effective incentives encourage organizations to consider risks, document decisions, and make proportionate investments. Ineffective incentives can lead to wasteful investments, or encourage the acceptance of risks where the consequences to the community greatly outweigh the consequences to the organization.

Sources:

Rick Nunes-Vaz, Steven Lord, and Daniel Bilusich, "From Strategic Security Risks to National Capability Priorities," *Security Challenges* 10, no. 3 (2014): 23–50.

Peter R. Orszag, "Critical Infrastructure Protection and the Private Sector: The Crucial Role of Incentives," *Journal of Risk and Uncertainty* 26 (2003): 231–49.

The following section will now describe the fundamental ways a national risk program can encourage and guide risk management practices to foster a mutually supportive environment for CISR.

Managing Security Risks

Too often, the treatment of risks is perceived to be a relatively straightforward process of addressing the identified risks in priority order. In reality, the analysis required to support risk management decisions is as complex as the analysis required to characterize and compare the risks in a coherent manner. A national risk program has a responsibility to assist local, regional, and national entities in managing risks, which can begin with characterizing and comparing them properly. Some risks are directly comparable, whereas others can be only indirectly compared, due to very different bases for considering threat, vulnerability, or consequences. Some risks are best considered in an immediate time frame, for example, while others require a longer-term perspective. Some threats are constant, in contrast to others that are rare or episodic. A national risk program can assist in the management of risks by organizing risk into portfolios.

Portfolios of risks contain directly comparable risks because they have similar considerations for risk factors. Natural hazards, for example,

can be meaningfully described through expected frequency of occurrence, structural vulnerability, community resilience, and a common set of consequences. Risks from terrorism, in contrast, usually consider relative threat, which involves understanding different terrorist groups, their opportunity, intent, and capability, which then heavily influences the assessment of vulnerability and consequence. Cyber threats, similarly, require a different manner of thinking about the likelihood that an organization or geographic region will encounter them. See chapter 3 for a discussion of infrastructure-specific intricacies of cybersecurity risks. Although there are some commonalities through all of these risks, it is best to define them within portfolios. Cross-portfolio analysis is relevant primarily in the evaluation of risk management solutions. That is, risk treatments or controls may affect risks in multiple portfolios, which helps decisionmakers account for the high-level comparisons among the portfolios without directly comparing individual risks across portfolios. See table 13-3 for an example of cross-portfolio analysis.

Table 13-3. Managing risks at the regional level

In 2008, the US National Capital Region (NCR) conducted a security risk assessment to examine the community's security risks. Working with the Office for National Capital Region Coordination within the Federal Emergency Management Agency, the NCR obtained input from federal, state, local, and private-sector partners in collaboration with the Department of Homeland Security's Office of Risk Management and Analysis. The risk analysis ranked potential critical infrastructure hazards by scenario and provided options for risk reduction.

The risk approach divided the risks into two portfolios, natural hazards and terrorism, allowing a direct comparison of risks with similar data availability. The visual comparison of risks within the portfolios enabled decision-makers not only to understand their relative severity, but also the risk factors that drove the severity. Beyond helping build a common understanding of the risks the NCR faced, the assessment helped leaders evaluate risk management investments. The analysis allowed the creation of strategic investment themes and identified the scenarios in both portfolios that would be affected. This process helped leaders understand how broad an investment's effects might be, as well as the extent to which it potentially reduced risk by addressing threat, vulnerabilities, or consequence.

Sources:

William O. Jenkins, Jr., *National Capital Region: 2010 Strategic Plan Is Generally Consistent with Characteristics of Effective Strategies* (Washington, DC: Government Accountability Office, 2011).

Elizabeth Jackson, William L. McGill, and Christopher Geldar, "Regional Risk Analysis: A Coordinated Effort" (presentation, Security and Risk Management Association Annual Conference, Arlington, VA, June 18, 2009).

To foster an environment in which each entity's efforts support and augment the efforts of related stakeholders, a national risk program should embrace two high-level approaches. First, it should help guide local efforts to maximize information exchange and compatibility of risk approaches so that higher level perspectives can be built from the bottom up. Second, it should ensure validity and accuracy when organizations communicate risk results to build a common understanding of risk. The following sections discuss these two approaches in more detail.

Building from the Bottom Up

Security risk assessments completed at the local level can benefit from the natural advantage of drawing on local experts who know the geography, hazards, and population well. At the local level, decisionmakers tend to be concerned primarily with the most immediate risks: those that occur frequently or those that are particularly pertinent to the area. This makes sense from the perspective that emergency managers want to maximize the investments that will bring the most assistance to addressing the most common problems. For a national risk program to benefit from the local expertise, it needs to standardize the risk analysis sufficiently to be able to compare different assessments but without being so rigid or complex that it draws resources away from emergency management efforts. To some extent, the most fundamental requirement is that the basis for considering threat, vulnerability, and consequence is sufficiently documented so external reviewers understand the criteria for evaluation and prioritization.

A national program may still provide guidance to help align these assessments by stipulating the time period for evaluation (for example, one to five years) or the scope (for example, the portfolios of risk or types of hazards considered). Where government agencies have regulatory or oversight authority over sectors of infrastructure, a national risk program can encourage information sharing with local governments to maximize a common understanding of the risks that infrastructure assets and systems have, while minimizing duplication of effort. Building a community that produces sound and comparable risk assessments enables the understanding of more complex risks, such as the dependency and interdependency analysis discussed in chapter 12. Local government may have its own requirements for how to evaluate risk management investments, but the national government can help make this consistent as well by stipulating the analyses of alternatives, including initial costs, out-year costs, opportunity

costs, the scenarios that benefit from the investment, and the degree to which the investments reduce risks.[8]

It is important that national risk programs enhance the relationship between various levels of government and delineate the responsibilities among them for assessing and managing security risks. Figure 13-1 is a conceptual model that depicts these relationships and responsibilities to assess and manage risk based on the likelihood of the risk occurrence and the severity of the consequence. It shows how local government is best suited to focus on likely and locally manageable events whereas international cooperation is required for the most probable, highest consequence events. State or regional governments augment and broaden local government risk management efforts while national-level government programs address the higher consequence events, even if they are low probability.

By ensuring local-level risk assessments are compatible and incorporating sector-specific information already collected, a national-level risk program can aggregate information to build regional and national-level perspectives.[9] Building this cross-jurisdiction comparison brings multiple advantages. First, it ensures due diligence for risk analysis at the local level. By comparing the risks that local jurisdictions identify and evaluate, the national program will gain a more comprehensive understanding of possible risks, encourage the inclusion of overlooked risks where appropriate, and share best practices, including potential data sources and analytic techniques. Second, the cross-jurisdictional perspective also enables gap analysis, identifying risks that are too strategic for local jurisdictions to consider (as figure 13-1 illustrates). Third, building a strong foundation for risk assessment in the near term allows the opportunity for an examination of longer-term risks. See table 13-4 for an example of one approach to assessing long-term risks. National risk programs should be able to think about investments more broadly, either over large geographic areas or across multiple risks.

8. Government Accountability Office (GAO), *Further Refinements Needed to Assess Risks and Prioritize Protective Measures at Ports and Other Critical Infrastructure* (Washington, DC: GAO, 2005), 106–7.

9. Michael H. Brody, "Enhancing the Organization of the United States Department of Homeland Security to Account for National Risk," *Homeland Security Affairs* 16 (2020): 21.

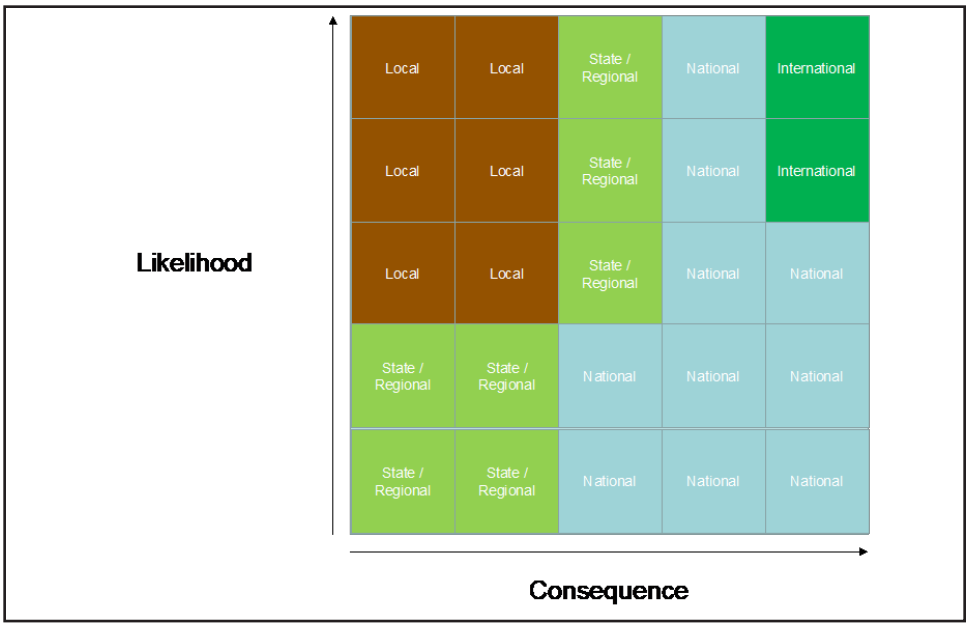


Figure 13-1. Levels of governmental responsibility for risk management

Table 13-4. Exploring low-likelihood, high-consequence risk scenarios

Considering risks that are in the future or that are extant but with low likelihood requires decisionmakers to cope with the high degree of uncertainty inherent in those scenarios. There are multiple ways of exploring these risks, one of which is an approach known as *alternative futures analysis*. This approach focuses on identifying two primary drivers of uncertainty and then exploring their interaction. The US National Intelligence Council has used this approach for its long-term analysis. A visual depiction of the two drivers of uncertainty creates a simple matrix, and exploring specific scenarios within each quadrant can make the risks more apparent.

It is important to understand the distinction between these structured analytic exercises and predictions. These types of analyses can offer many benefits to leadership, but not if they limit understanding and action to focus on a few specific scenarios. Instead, this approach should allow leadership to explore the longest-term effects that the emerging trends may have, envision the outcomes or innovations most beneficial to homeland or national security, and consider the partners required to shape those risks. Leaders can then examine certain aspects of emerging risks more closely and identify policies, guidance, or other tools to maximize resilience in the critical infrastructure community.

Sources:

Central Intelligence Agency (CIA), Structured Analytic Techniques for Improving Intelligence Analysis (Washington, DC: CIA Center for the Study of Intelligence, 2009).

National Intelligence Council, *Mapping the Global Future: Report of the National Intelligence Council's 2020 Project* (Washington, DC: Government Printing Office, 2004).

Neither threats nor infrastructure systems are bound to national borders. Transportation, energy, and telecommunications have inescapable international dimensions that must be addressed. Without cross-border coordination, attempts to secure a single part of an international system risk being wasted effort. Work with allies can be straightforward if not simple. Agreement on policy and standards can bridge many of the potential gaps in security. Information sharing and reciprocity, and even shared technology, can minimize the friction in some systems as they cross borders. See chapter 10 for a discussion of how the European Union has addressed this aspect of risk management.

Countries must work with their allies but also with their competitors and adversaries. In these cases, significant constraints will impede sharing some technologies, and potentially even security standards, but the basic foundation is the same. Information sharing and communication that establish a shared understanding of the benefits of secure, resilient infrastructure can open pathways to parallel programs or even shared investments that secure transportation modalities, supply chains, and access to energy and telecommunications systems, among others.

Building a Common Understanding of Risks

As previously mentioned, a national risk program has a responsibility to coordinate and integrate risk analysis among the myriad public and private-sector partners involved in CISR. This coordination does not mean forcing a one-size-fits-all approach, but it does mean encouraging a common vocabulary for security risk professionals so the constituent parts of the community can communicate readily with each other. It also means creating mechanisms or rubrics for making risk analyses comparable where they cannot be uniform.

There are numerous ways to define security risk, its factors, and their combinations.[10] Several factors influence what approach any organization will adopt, especially the organization's capabilities and level of organizational maturity. External factors include whether data are available and whether acceptable quantitative approaches exist. For security risk, it can be considerably challenging to quantify threat, vulnerability, and consequence, and qualitative or semi-quantitative approaches are the only approaches to assessing emerging risks. The COVID-19 pandemic is an example of a situation in which decisions required security risk

10. Julian Talbot and Miles Jakeman, *Security Risk Management Body of Knowledge* (Hoboken, NJ: John Wiley & Sons, 2011), 142–47.

analysis before data were available, but qualitative approaches could provide insight.[11]

Quantitative approaches are best suited for organizations with a high level of maturity in applying risk to decision making, and where data exist to compare scenarios that are sufficiently similar. When combining or comparing risk analyses from multiple organizations or for different risks, quantitative approaches may be difficult to create or potentially not possible.[12] National risk programs, therefore, must promulgate best practices in the CISR field to maximize comparability, create ways to meaningfully compare and combine individual risk assessments, and clarify when to avoid such integration. By helping stakeholders and leaders understand the relative severity of risks, the ways they differ, and which risks are categorically different, national risk programs can build a common understanding of risk and promote educated discussions on how to manage them.

Key to this common understanding is the reality that the majority of security risks cannot be eliminated but only managed, and that management requires a discussion of trade-offs. For example, airport security can be raised significantly, at great cost and at reduced passenger throughput, but societies are unlikely to tolerate such conditions for long. See chapter 6 for detail on the trade-offs involved in aviation security and chapter 7 for a discussion of how the multifaceted nature of railways complicates security investment decisions. In any operational environment, risk controls bring an initial cost, as well as opportunity costs, which must be addressed directly. For critical infrastructure environments especially, the risk control's effect on the functioning of the infrastructure asset or system is also a necessary consideration when identifying, comparing, and selecting alternative investments. See chapter 8 for an example of such risk management decisions for the water sector. A national risk program can help the CISR community implement concepts (such as managing risk to "as low as reasonably practicable") or other standards that help stakeholders understand their individual and shared responsibilities.

11. See Rachael J. Oakenfull and Anthony J. Wilson, *Qualitative Risk Assessment: What is the Risk of Food or Food Contact Materials Being a Source or Transmission Route of SARS-CoV-2 for UK Consumers* (London: Food Standards Agency, 2020).

12. Peter Månsson, "Uncommon Sense: A Review of Challenges and Opportunities for Aggregating Disaster Risk Information," *International Journal of Disaster Risk Reduction* 40 (2019): 6–9.

Necessary Characteristics of High-quality Risk Programs

Risk analysis and risk management approaches can accommodate different levels of maturity and quantification. An entirely qualitative approach to risk, for instance, can be both useful and credible, especially with issues where key data or a valid method of quantification are not available.[13] There are three key aspects of a risk program, however, that are not negotiable. As NATO member states and partner nations consider and update their national risk programs, this concluding section will highlight the need to include these three elements: the transparency of the decision-making process, clear risk communication, and coherent risk governance.

Transparency

Managing risk affects a large number of stakeholders. While the benefits should convey to a large number of people, the costs may be unevenly distributed. Not only are there opportunity costs—the investments in people, equipment, or training that could have been used in other areas—but some stakeholders may be exposed to higher risks for the overall management of a greater risk. The construction of a biosafety laboratory, for example, may help the nation fight dangerous pathogens in the long term but it may not be welcome in a community in the short term. Building an additional power plant may strengthen the electric power grid but its construction also brings environmental or other concerns to the communities nearby. Incomplete or poorly documented risk assessments preclude any external entities—including people affected by the risk-based decisions—from understanding the risk analysis, invalidating the analysis regardless of the quality of the conclusions.[14]

If the communities are to support the risk management decisions, then leaders need to involve stakeholders at several key steps. This involvement includes the generation of risk scenarios, the evaluation of consequences, the evaluation of risk treatments, and the monitoring of effectiveness. If communities understand the basis for the decisions, they are much more likely to accept the decisions. Where communities have been excluded from these stages of risk management, they have sometimes balked at the final decisions, leading to delays or cancellation of the risk controls.

13. National Research Council, *Understanding Risk: Informing Decisions in a Democratic Society* (Washington, DC: National Academies Press, 1996), 97.

14. National Research Council, *Technical Input on the National Institutes of Health's Draft Supplementary Risk Assessments and Site Suitability Analyses for the National Emerging Infectious Diseases Laboratory, Boston University* (Washington, DC: National Academies Press, 2007), 11–12.

Building in opportunities for appropriate engagement with stakeholders improves the quality of the processes and makes the outcomes more likely to be accepted.[15]

Risk Communication

Communicating the results of risk analysis can be particularly challenging. Conveying the findings in terms that the widest range of stakeholders can understand often requires moving away from the scientific or technical language that analysts use, and adopting plain language. The challenge is that risk analysis often has nuanced conclusions, and stating the findings too simply may over- or understate the likelihood of the risk or the severity of the consequences. A failure of risk communication can cause panic or complacency. Situations where political sensitivities overrule scientific or professional consensus undermine the public's understanding of risks, which can lead to catastrophic consequences.[16] Risk communication must begin prior to the emergence of a specific threat or crisis; building relationships prior to a crisis increases trust and can increase a community's risk tolerance level.[17]

Organizations must plan to communicate to the public through multiple methods to get specific messages out effectively to appropriate audiences, especially during times of crisis. See chapter 15 for an in-depth discussion of crisis management capabilities and tools. When crises do arrive, risk advisory organizations must ensure timely notification of risks and provide specific information on the nature, location, and timing of threats, as well as guidance on actions to take in response to threats.[18] When risk communications are late, vague, or perceived to be arbitrary, they erode public confidence and increase the barriers for future communication efforts.[19]

15. Bruce Altevogt, Megan Reeve, and Theresa Wizemann, eds., *Engaging the Public in Critical Disaster Planning and Decision Making: Workshop Summary* (Washington, DC: National Academies Press, 2013).

16. Stephen S. Hall, "Scientists on Trial: At Fault?," *Nature* 477, no. 7364 (2011): 264–69.

17. Melissa Janoske, Brooke Liu, and Ben Sheppard, *Understanding Risk Communication Best Practices: A Guide For Emergency Managers and Communicators* (College Park: National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland, 2012), 20–21.

18. GAO, *Homeland Security: Communication Protocols and Risk Communication Principles Can Assist in Refining the Advisory System* (Washington, DC: GAO, 2004), 12–13.

19. Jacob N. Shapiro and Dara Kay Cohen, "Color Blind: Lessons from the Failed Homeland Security Advisory System," *International Security* 32, no. 2 (2007): 121–54.

Risk Governance

As mentioned above, most nations will have a number of different organizations contributing to the understanding of security risks, and a national risk program should connect them into a coherent network. Part of that effort includes clearly assigning roles and responsibilities to local or regional government organizations and other partners in the private and public sectors. Ensuring that each organization has a clear understanding of which risks it should take into account helps build a collaborative environment where individual organizations see their roles in relation to other partners.[20] This connection fosters a culture of accountability and encourages organizations to raise risks where there are shared responsibilities. Regular reports to oversight offices will help senior leaders understand which risks are being addressed (as well as which need additional action) and maximize the probability that risks where there is not clear authority or responsibility can be identified and directed. When national risk programs allow high degrees of ambiguity, they minimize the likelihood senior leaders will recognize situations where responsibilities overlap but suffer from a lack of collaboration or where significant risks fall between the authorities of multiple agencies. Clear, coherent, strategic-level assessments of risks in each major portfolio can be one of the primary mechanisms that allow national governments to understand the security risks they face and whether risk management actions are sufficient.[21]

20. GAO, *Enterprise Risk Management: Selected Agencies' Experiences Illustrate Good Practices in Managing Risk* (Washington, DC: GAO, 2016), 6.

21. Thomas Cooper, *Strategic Risk Management in the Municipal and Public Sector: An Exploration of Critical Success Factors and Barriers to Strategic Risk Management within the Province of Newfoundland and Labrador* (St. John's, Canada: Memorial University, 2010), 69–75.

– 14 –

Enhancing Cybersecurity of Industrial Control Systems

Sungbaek Cho

Due to the advancement of information and communications technologies, most modern critical infrastructure operates electronically. Malevolent forces could exploit any weaknesses or vulnerabilities in the devices and equipment that comprise these critical infrastructure systems to launch cyberattacks that adversely affect the society and its national security. For instance, cyber incidents targeting lifeline sectors—such as electricity, water supply, and transportation—may not simply lead to inconvenience and financial losses for people and businesses, they can also cause social turmoil, disruption of military operations, and human casualties or fatalities. For these reasons, most countries regard the cyber defense of critical infrastructure systems and assets as a top priority, and they are undertaking extensive efforts to enhance their critical infrastructure security and resilience (CISR) posture.

The North Atlantic Treaty Organization identifies cyberattacks against critical infrastructure as a possible instability situation, defined as a future event significant enough to reach the threshold requiring the Alliance to use military forces.[1] As national and societal functions rely heavily on information technology, improving cybersecurity has become a significant element of member states' efforts to enhance national CISR. Similarly, NATO has identified the important link between cybersecurity and the Alliance's

1. NATO, *Framework for Future Alliance Operations: 2018 Report* (Norfolk, VA: Allied Command Transformation, 2018), 15, https://www.act.nato.int/images/stories/media/doclibrary/180514_ffao18.pdf.

ability to fulfill its core tasks. At the Warsaw Summit in 2016, NATO officially recognized cyberspace as a domain of operations in which the Alliance must "defend itself as effectively as it does in the air, on land, and at sea."[2] At Warsaw, the Allies also pledged to strengthen and enhance the cyber defenses of national networks and critical infrastructure as a matter of priority, highlighting that NATO as an organization is only as strong as its weakest link.[3] NATO now serves as a venue in which Allies can consult on cyber defense issues, share information on cyber threats, exchange best practices, and coordinate activities including education, training, and exercises.[4]

Depending on its scale and severity, a cyberattack against a NATO member state's critical infrastructure could be regarded in the same way as an armed attack that would justify the targeted country's right to self-defense.[5] A destructive cyberattack also could lead Allies to invoke Article 5 of the Washington Treaty—the mutual defense clause that states an attack against one Ally is an attack against all Allies—though such a decision would be taken by the North Atlantic Council on a case-by-case basis.[6] In response to the evolving cyber threat landscape, NATO's stance against cyberattacks was further extended at the Brussels Summit in 2021, where Allied leaders recognized that the impact of cumulative, malicious cyber activities could amount to an armed attack.[7] The term *cumulative* implies several lower-impact cyberattacks by the same adversary over time could be as destructive as a single, massive cyberattack.[8] Regarding cyber operations against adversaries, NATO doctrine introduces a concept known as Sovereign Cyber Effects Provided Voluntarily by Allies, a mechanism that allows individual member states to support voluntarily other Allies' offensive cyber capabilities in the case of armed

2. "Warsaw Summit Communiqué," NATO (website), July 9, 2016, https://www.nato.int/cps/en/natohq/official_texts_133169.htm.

3. "Cyber Defence Pledge," NATO (website), July 8, 2016, https://www.nato.int/cps/su/natohq/official_texts_133177.htm.

4. "Fact Sheet: NATO Cyber Defence," NATO (website), August 2020, https://www.nato.int/nato_static_fl2014/assets/pdf/2020/8/pdf/2008-factsheet-cyber-defence-en.pdf.

5. Michael N. Schmitt, general ed., *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations,* 2nd ed., managing ed. Liis Vihul (Cambridge, UK: Cambridge University Press, 2017), 339–48.

6. "Wales Summit Declaration," NATO (website), September 5, 2014, https://www.nato.int/cps/en/natohq/official_texts_112964.htm.

7. "Brussels Summit Communiqué," NATO (website), June 24, 2021, https://www.nato.int/cps/en/natohq/news_185000.htm.

8. Stefan Soesanto, "When Does a 'Cyber Attack' Demand Retaliation? NATO Broadens Its View," Defense One (website), June 30, 2021, https://www.defenseone.com/ideas/2021/06/when-does-cyber-attack-demand-retaliation-nato-broadens-its-view/175028/.

conflicts, and outlines the procedures for defensive cyber operations, including self-defense and collective defense.[9]

Although NATO is taking steps to improve its collective ability to defend against and respond to cyberattacks against Allied critical infrastructure, it should be kept in mind that individual member states form the first line of defense. Thus, enhancing cyber defense capabilities and enhancing CISR policies and procedures are the primary responsibilities of each Ally. With these objectives in mind, this chapter aims to provide an overview of the major cybersecurity issues surrounding critical infrastructure with a special focus on industrial control systems (ICS). Based on this understanding, the chapter will offer best practices and tools for critical infrastructure stakeholders, owners, and operators to protect their systems and enhance security and resilience against cyberattacks.

An Overview of Industrial Control Systems (ICS)

To understand cybersecurity requires a proper knowledge of ICS. The term ICS includes various control systems typically found in industrial sectors and critical infrastructure. Also known as operational technology (OT), an ICS consists of combinations of different control components (electrical, mechanical, hydraulic, and pneumatic, for instance) to achieve an industrial objective, such as manufacturing, transportation, or energy.[10] Examples of ICS include power plants, electrical grids, water and water treatment systems, energy transport, and railways. While an ICS can be configured and operated in a variety of ways, there are three common control systems that merit further explanation.[11]

- Supervisory control and data acquisition (SCADA) systems are used to control dispersed assets centrally. Typical examples are water distribution, wastewater collection, power grids, railways, and other public transportation systems.

9. NATO Standardization Office, *Allied Joint Doctrine for Cyberspace Operations,* Allied Joint Publication 3.20 (Brussels: NATO, 2020), 5, https://www.gov.uk/government/publications/allied-joint-doctrine-for-cyberspace-operations-ajp-320.

10. National Institute of Standards and Technology (NIST), *Guide to Industrial Control Systems (ICS) Security* (Washington, DC: Department of Commerce, 2015), B-8, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r2.pdf.

11. NIST, *Industrial Control Systems*, 2-5–2-13.

- Distributed control systems (DCS) manage continuous production processes within the same geographic area. Examples include oil refineries, water and wastewater treatment facilities, power plants, chemical plants, and pharmaceutical processing facilities.
- Programmable logic controllers (PLC) are devices that control discrete processes, such as automobile assembly lines. While a PLC is often used as a component for a SCADA system or DCS, it can also be implemented as the primary controller in a small ICS.

While actual ICS architectures vary widely based on the nature of the critical infrastructure sector and type of facility, the Purdue reference architecture is widely recognized as the standard model for common control systems.[12] Having a model that depicts the control system architecture and shows the various interconnections between technological components can help organizations segment the various networks, develop zones with clear boundaries, and create layers of cyber defense measures. The US Department of Homeland Security (DHS) recommends this process of developing a secure network architecture as a means to limit cyber threat actors' ability to exploit ICS, which is far easier when the systems are integrated and no zones or boundaries exist.[13] The Department of Homeland Security endorses developing a layered cyber defense consisting of five unique zones, as outlined in figure 14-1.[14]

12. Theodore Williams, "The Purdue Enterprise Reference Architecture," *IFAC Proceedings* 26, no. 2 (1993): 559–64, https://doi.org/10.1016/S1474-6670(17)48532-6.

13. Department of Homeland Security (DHS), *Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies* (Washington, DC: DHS, 2016), 16–20, https://us-cert.cisa.gov/sites/default/files/recommended_practices/NCCIC_ICS-CERT_Defense_in_Depth_2016_S508C.pdf.

14. DHS, *Recommended Practice*, 18.

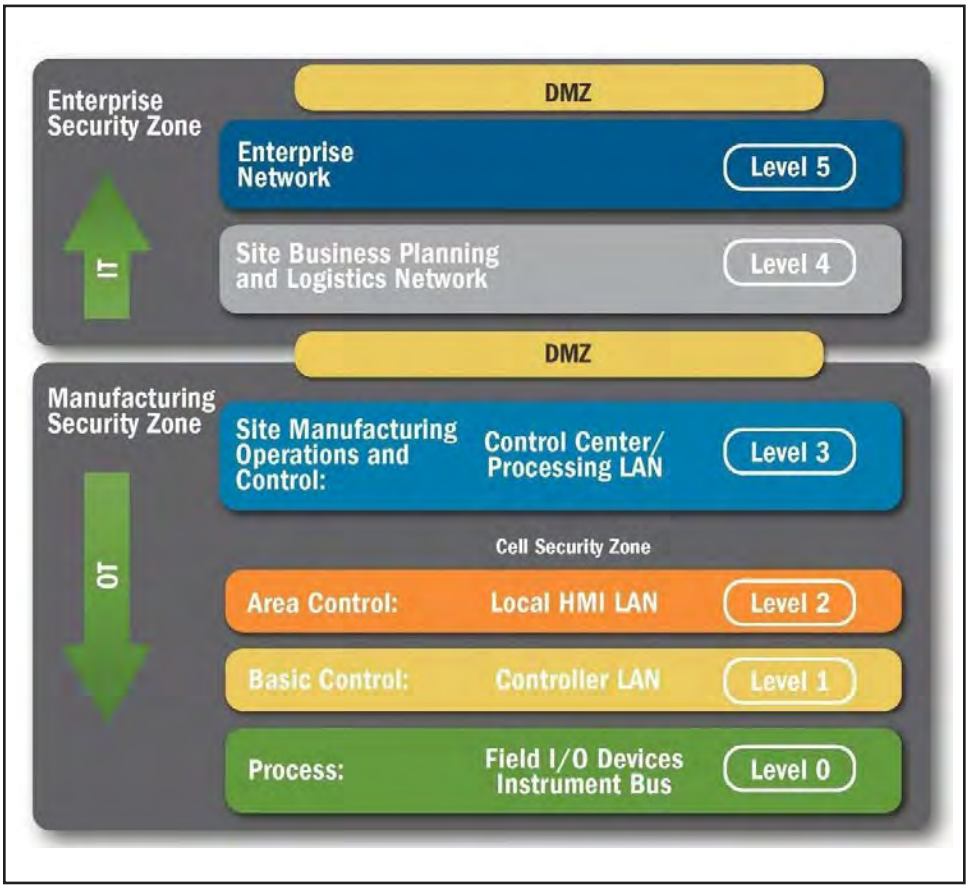


Figure 14-1. ICS reference model
(Diagram by US Department of Homeland Security)

The first section, the enterprise security zone, is not directly related to the ICS. This zone provides employees the connectivity to the Internet, remote sites, and business networks that comprise the intranet, e-mail servers, web servers, and other business systems. The enterprise security zone is also known as the informational technology (IT) system. See chapter 3 for its helpful explanation of operational and informational technology [OT and IT] systems. Next, the manufacturing security zone is where a vast majority of monitoring and control takes place. Depending on the size of the ICS, this zone may contain multiple cell zones. The third section, the cell zone, contains local human-machine interfaces (HMI), controllers, and field devices to be monitored and controlled. The HMI is a desktop computer with control software through which operating personnel manipulate the ICS. The cell zone also may include a safety instrumented system, which is a special controller designed to automatically take actions in the event

of dangerous conditions like excessive pressure or temperature. Fieldbus protocols with hard wiring are typically used between field devices and controllers, whereas Ethernet is common between controllers and HMIs. Finally, a demilitarized zone (DMZ) is a subnetwork that acts as an intermediary to protect the inside network. Within the ICS, the DMZ is where the data historian, antivirus or patch, and remote access gateway are located. A data historian is a time-series database to capture all production and process data for monitoring and analysis troubleshooting.

Security Concerns in ICS

In the past, critical infrastructure facilities operated ICS strictly in a closed network environment. To ensure real-time monitoring and efficient and effective resource planning at the enterprise level, however, the prevalent practice in modern critical infrastructure is to operate ICS in a more open, interconnected network with business networks. Examples of business applications that may connect to ICS include production planning and scheduling applications, manufacturing execution systems, inventory management systems, and maintenance management systems.[15] Furthermore, the Ethernet and other open standard technologies are also becoming more prevalent in ICS. As a result, attackers can understand and exploit system components more easily than they could in the past. These realities raise security concerns because ICS are more vulnerable to cyberattacks than ever before. When compared to IT systems, the following system characteristics make it more challenging to secure ICS in the face of the numerous vulnerabilities, risks, and threats in the cyber domain.[16]

- Timeliness and performance requirements. As ICS are usually time-critical, security measures causing an unacceptable delay and/or threatening the functionality of the system cannot be deployed.
- Availability requirements. Patches cannot be applied on time as they have to be tested thoroughly for stability and reliability. Outages of systems to install patches typically must be scheduled weeks in advance.
- Risk management requirements. Security measures that impair safety are unacceptable.

15. Eric D. Knapp and Joel T. Langill, *Industrial Network Security* (Waltham, MA: Syngress, 2014), 20.
16. NIST, *Guide to Industrial Control Systems* (ICS) (2015), 2-14–2-17.

- Physical effects. As ICS have complicated physical processes, good communications between experts in the control system and the physical domain are necessary.
- System operation. Since ICS operating systems and networks are often quite different from IT counterparts, they require different skill sets, experience, and levels of expertise.
- Resource constraints. Many components are resource-constrained in memory and processing power. As a result, typical contemporary security capabilities may not apply.
- Communications. Communication protocols and media for field devices (sensors and actuators) are different from those used in IT environments and thus require other specialties.
- Managed support. Given the fact that maintenance is often performed by a single vendor, the use of third-party solutions requires the vendor's approval or the ICS will no longer be under warranty.
- Component lifetime. The lifetime of the ICS components is often over 15 years, while IT components require upgrades and patches much more frequently.
- Component location. In some cases, ICS components may be located at remote sites that require extensive transportation effort to reach. Each site needs to be appropriately protected.

Beyond the limitations and restrictions in applying sufficient security measures due to the inherent nature of ICS, there are also several security concerns and problems commonly found in most ICS.

Vulnerabilities in ICS Components

According to a recent cybersecurity survey, organizations disclosed 893 vulnerabilities specific to their ICS in 2020—a steady increase from the 672 reported in 2018 and the 716 in 2019.[17] Surprisingly, in 76 percent of these disclosed vulnerabilities, threat actors were able to launch attacks without needing to be authenticated. These figures, however, do not include vulnerabilities found in common IT components, such as employees' personal desktops, servers, databases, and network switches.

17. Claroty Research Team, *Claroty Biannual ICS Risk & Vulnerability Report: 2H 2020* (New York: Claroty, 2020), 4–11, https://security.claroty.com/biannual-ics-risk-vulnerability-report-2H-2020.

These components are predominantly commercial-off-the-shelf products or custom-made models based on these products. Traditionally, vendors have not considered security an integral part of a product development process, but this dynamic is changing. Despite the recent rise in concern regarding security of control system components during product development, the level of security in ICS lags behind and is not as comprehensive when compared to the security of IT products.[18] Therefore, there are many weaknesses in ICS components, including susceptibility to denial-of-service attacks and lack of security checks for firmware updates. Even IT components used in control systems are often configured to enable insecure services, such as Telnet, by default.[19]

ICS Components Exposed to the Internet

Many ICS components are connected to the Internet without proper security measures like firewalls or remote access gateways. In 2019, a search on Shodan—a special search engine often used to find devices connected to the Internet—revealed more than 2.6 million ICS components around the globe were connected to the Internet.[20] Most of these devices were likely used in schools for research or by small private companies. Poor security practices or breaches in security protocols (such as opening a firewall port for remote access and then forgetting to close it or connecting to the Internet intentionally to reduce work burdens) may occur even in national critical infrastructure, making these facilities and organizations equally vulnerable.

Connection with Business Systems

According to the SANS Institute's 2019 survey of 338 organizations, 57 percent connected their ICS to business systems while 35 percent connected their ICS to the Internet either through the DMZ or directly.[21] When such a connection is inevitable, it must be secured to prevent malicious traffic from entering the ICS network. A firewall can be used for this purpose, but a unidirectional network device—a special security gateway that is also

18. DHS, *Recommended Practice*, 4.

19. Joseph Weiss, *Protecting Industrial Control Systems from Electronic Threats* (New York: Momentum Press, 2010), 29.

20. David Hasselquist, Abhimanyu Rawat, and Andrei Gurtov, "Trends and Detection Avoidance of Internet-Connected Industrial Control Systems," *IEEE Access* 7 (2019): 155504–12, https://doi.org/10.1109/ACCESS.2019.2948793.

21. Barbara Filkins, *SANS 2019 State of OT/ICS Cybersecurity Survey* (Rockville, MD: SANS Institute, 2019), 12.

known as a data diode—is the optimal solution because it is allows data to travel in only one direction.[22] Organizations may also consider using an intrusion detection system (IDS). If such security devices are in place, however, there is still a risk of allowing malicious traffic due to misconfiguration. Moreover, an IDS cannot be used if the control system vendor does not approve because of potential degradation in network performance. An IDS is more commonly found in IT networks than in ICS networks and, even when installed, it may not be able to fully understand ICS protocols.[23]

Outdated Components

As an ICS typically has a very long lifespan, it is common to find ICS components already past end of life, such as when HMIs run on outdated programs like Windows XP or 7. Even if organizations want to upgrade old components, they cannot be upgraded if application software does not support the latest operating system or the vendor does not guarantee reliability after an upgrade. Installing antivirus programs on old desktops may not be feasible because of performance and stability issues. Moreover, when old hardware is damaged, it may not be easy to find replacement options that meet the same specifications.

Remote Access to Control Networks

With the recent development of cloud technology, cloud-based management services for IT systems have emerged, and similar movements are also emerging for ICS. According to the 2019 SANS survey previously cited, more than 40 percent of respondents used cloud-based services for their ICS. Respondents gave three main reasons for why they used these services: (1) remote monitoring, (2) configuration, and (3) analysis, which accounted for 44 percent of the reported uses.[24] Regardless of the types of outsourced services, all remote accesses must be controlled in a highly secure manner.

Insecure Nature of ICS Protocols

All major fieldbus protocols—such as Modbus, DNP3, Profinet, and EtherCAT—are susceptible to man-in-the-middle attacks because they generally lack sufficient authentication or encryption.[25] Such attacks can disrupt network operations or manipulate input-output messages

22. NIST, *Industrial Control Systems*, 5–21.

23. European Union Agency for Cybersecurity (ENISA), *Communication Network Dependencies for ICS/SCADA Systems* (Athens: ENISA, 2016), 30, https://www.enisa.europa.eu/publications/ics-scada-dependencies/.

24. Filkins, *OT/ICS Cybersecurity Survey*, 13–14.

25. Knapp and Langill, *Industrial Network Security*, 166.

to cause failure. Protocol gateways, including serial-to-Ethernet converters, that translate one ICS protocol to another could provide an additional attack vector as they may contain security flaws and vulnerabilities.[26]

Major Cyber Incidents

Due to insecure configuration and management, cyber incidents in ICS have unfortunately become a common occurrence. This section will now examine some of the significant cyberattacks that targeted ICS.

Stuxnet (2010)

The most historic cyber incident associated with ICS was the infection of Iran's nuclear program with Stuxnet, a worm designed to penetrate air-gapped control networks via USB flash drives and then propagate through self-replication. The Stuxnet worm, which was discovered in 2010, precisely targeted the centrifuges used in Iran's uranium enrichment process to change the frequencies of the frequency converters covertly that adjust motor speed. It is activated only when the same software—namely, Siemens WinCC and Step7—and frequency range as the Iranian facility are found.[27] While the physical consequences of Stuxnet were limited in that Iran took just one year to recover fully from the effects of the attack, this incident demonstrated that separating the ICS network from the Internet can no longer be considered a sufficient security measure.[28]

BlackEnergy (2011)

In 2014, the US Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) alerted that BlackEnergy malware had been targeting users of HMI products, such as GE Cimplicity, Advantech/Broadwin WebAccess, and Siemens WinCC, since 2011.[29] Attackers targeted the Internet-connected HMIs and then exploited a vulnerability of the software to install BlackEnergy

26. Marco Balduzzi et al., *Lost in Translation: When Industrial Protocol Translation Goes Wrong* (Irving, TX: Trend Micro, 2020), 48–49, https://i.blackhat.com/USA-20/Wednesday/us-20-Balduzzi-Industrial-Protocol-Gateways-Under-Analysis-wp.pdf.

27. William Maclean, "Stuxnet Study Suggests Iran Enrichment Aim: Experts," Reuters (website), November, 16, 2010, https://www.reuters.com/article/us-security-cyber-stuxnet-idUSTRE6AF2F320101116.

28. Marie Baezner and Patrice Robin, "CSS Cyber Defense Hotspot Analysis Issue 4: Hotspot Analysis: Stuxnet," Center for Security Studies at ETH Zurich (website), October 2017, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Cyber-Reports-2017-04.pdf.

29. "ICS Alert: Ongoing Sophisticated Malware Campaign Compromising ICS (Update E)," Cybersecurity and Infrastructure Security Agency (CISA) (website), July 22, 2021, https://us-cert.cisa.gov/ics/alerts/ICS-ALERT-14-281-01B.

malware. Although no malicious activity was identified, the malware could have damaged, modified, or disrupted the targeted systems. A security company found that some of the command and control (C2) servers used in this attack were the same as those used by the Russian Advanced Persistent Threat (APT) group known as Sandworm.[30] In July 2021, the US government officially attributed the BlackEnergy attack to Russian nation-state cyber actors.[31]

Havex (2013)

The Russian APT group known as Dragonfly used Havex in a cyber espionage campaign targeting ICS in a variety of countries, including several NATO member states.[32] Havex is a remote access Trojan that leveraged the Open Platform Communications—the data exchange protocol between Windows systems and controllers—to collect information on the targeted devices. The attackers Trojanized software available for download from three ICS manufacturer websites and gained access to the networks of systems that had installed the software.[33] A security company later found 88 variants were communicating with 146 C2 servers, which made connections with 1,500 different Internet Protocol addresses, each of which represents a possible victim of the attack.[34] Although the primary usage of Havex was espionage, its C2 server could have also been used in other attacks.[35] In 2021, the US government attributed the Havex attacks to Russia.[36]

German Steel Mill (2014)

According to the annual report issued in 2014 by Germany's Federal Office for Information Security (*Bundesamt für Sicherheit in der Informationstechnik*, or BSI), unspecified threat actors attacked a German steel mill, compromising individual ICS components and causing a furnace to shut down

30. Kyle Wilhoit and Jim Gogolinski, "Sandworm to Blacken: The SCADA Connection," *Trend Micro* (blog), October 16, 2014, https://blog.trendmicro.com/trendlabs-security-intelligence/sandworm-to-blacken-the-scada-connection/.

31. "Ongoing Sophisticated Malware Campaign."

32. Symantec, *Dragonfly: Cyberespionage Attacks against Energy Suppliers* (Mountain View, CA: Symantec, 2014), 5, https://docs.broadcom.com/doc/dragonfly_threat_against_western_energy_suppliers.

33. "ICS Advisory (ICSA-14-178-01): ICS Focused Malware," CISA (website),updated on July 20, 2021, https://us-cert.cisa.gov/ics/advisories/ICSA-14-178-01.

34. Daavid Hentunen and Antti Tikkanen, "Havex Hunts for ICS/SCADA Systems," F-Secure Labs (website), June 23, 2014, https://archive.f-secure.com/weblog/archives/00002718.html.

35. "ICS Alert (IR-ALERT-H-16-056-01): Cyber-Attack against Ukrainian Critical Infrastructure," CISA (website), updated July 20, 2021, https://us-cert.cisa.gov/ics/alerts/IR-ALERT-H-16-056-01.

36. "ICS Focused Malware."

in an abnormal manner.[37] The attackers used spear-phishing e-mails to steal login credentials and then used them to gain access to the mill's control system.

Ukraine Blackout (2015)

The Ukraine blackout in December 2015, which caused electricity disruption to 225,000 people in western Ukraine for up to six hours, was the first known successful cyber intrusion to take a power grid offline and one of the most severe incidents in cybersecurity history. The attackers, part of the Sandworm group, conducted a remote intrusion into three power distribution companies.[38] The attackers reportedly used spear phishing to obtain credentials in advance, which enabled the intrusion into the companies and then to the various substations.[39] Moreover, they infected Windows systems with KillDisk malware to erase files and the master boot record and corrupted the firmware of serial-to-Ethernet converters at substations to make them inoperable. As with the BlackEnergy and Havex attacks, the US government also attributed the 2015 blackout to Russia.[40]

RWE's Nuclear Power Plant, Germany (2016)

Computer viruses Conficker and W32.Ramnit were discovered in German utility company RWE's nuclear power plant near Munich in April 2016. The infected system was a computer used to view the movement of nuclear fuel rods, but the infection did not cause any harm as the plant was disconnected from the Internet.[41] The same malware was found on 18 removable drives used for office computers, implying that at least one of the office drives was inserted into the infected system. The official investigation also concluded the malware probably came from a USB drive.[42]

37. *Bundesamt für Sicherheit in der Informationstechnik* (BSI), *Millionenfacher Identitätsdiebstahl in Deutschland* (Bonn: BSI, 2014), https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2014.pdf.

38. Michael Assante, "Confirmation of a Coordinated Attack on the Ukrainian Power Grid," SANS Institute (website), January 6, 2016, https://www.sans.org/blog/confirmation-of-a-coordinated-attack-on-the-ukrainian-power-grid/.

39. "Cyber-Attack against Ukrainian Critical Infrastructure."

40. "Cyber-Attack against Ukrainian Critical Infrastructure."

41. Christoph Steitz and Eric Auchard, "German Nuclear Plant Infected with Computer Viruses, Operator Says," Reuters (website), April 26, 2016, https://www.reuters.com/article/us-nuclearpower-cyber-germany/german-nuclear-plant-infected-with-computer-viruses-operator-says-idUSKCN0XN2OS.

42. "Virus in the Gundremmingen Nuclear Power Plant Came from a USB Stick," CIO (website), June 3, 2016, https://www.cio.de/a/amp/virus-im-akw-gundremmingen-kam-ueber-usb-stick,3229370.

CrashOverride (2016)

During its cyberattack against a Ukrainian substation in December 2016 that caused a small-scale power outage, the Sandworm group used CrashOverride malware (also known as Industroyer).[43] This attack, like the several of the previous examples, was later attributed to Russian nation-state cyber actors.[44] Although this cyberattack was smaller in scale and duration than the one that caused the Ukraine blackout, CrashOverride was developed to create a far more widespread outage than the one that occurred in 2015. The CrashOverride malware has capabilities to issue malicious commands directly to remote terminal units—the controllers used for SCADA systems (such as power grids)—by exploiting the lack of authentication and authorization in the ICS protocol. The malware can also prevent legitimate communications with field devices, cause the shutdown of a relay, and employ its wiper module to render windows system inert and thus require a rebuild or backup restoration.[45] After the Stuxnet attack, the use of CrashOverride malware in 2016 is only the second known case of malicious codes intentionally built to disrupt physical systems. For a more detailed explanation and assessment of cyberattacks on Ukraine's power grid, see the overview provided in chapter 5.

TRITON (2017)

Following the mysterious shutdown of an entire petrochemical plant in Saudi Arabia in 2017, the subsequent investigation found the attackers gained remote access to an engineering workstation—a computer used for configuring a safety instrumented system (SIS)—using TRITON malware. TRITON, also known as TRISIS, is a malware that attacks the Triconex SIS fabricated by the company Schneider Electric. The TRITON malware allowed the attackers to reprogram the SIS, causing the controllers to shut down automatically.[46] Although it is not certain who is responsible for the cyberattack, evidence suggests Russia's Central Scientific Research Institute of Chemistry and Mechanics supported the development of TRITON.[47] In October 2020,

43. Assante, "Confirmation of a Coordinated Attack."

44. "Alert (TA17-163A): CrashOverride Malware," CISA (website), updated on July 20, 2021, https://us-cert.cisa.gov/ncas/alerts/TA17-163A.

45. "Alert: CrashOverride Malware."

46. Blake Johnson et al., "Attackers Deploy New ICS Attack Framework 'TRITON' and Cause Operational Disruption to Critical Infrastructure," *FireEye* (blog), December 14, 2017, https://www.fireeye.com/blog/threat-research/2017/12/attackers-deploy-new-ics-attack-framework-triton.html.

47. FireEye Intelligence, "TRITON Attribution: Russian Government-Owned Lab Most Likely Built Custom Intrusion Tools for TRITON Attackers," Mandiant (website), October 23, 2018, https://www.mandiant.com/resources/triton-attribution-russian-government-owned-lab-most-likely-built-tools.

the US Department of the Treasury imposed sanctions on this Russian research institution for its involvement with TRITON.[48]

Water Treatment Plant, United States (2021)

In February 2021, an unidentified attacker hacked the water treatment plant in Oldsmar, Florida. After accessing the plant remotely, the attacker tried to increase the level of sodium hydroxide in the water supply to 100 times greater than normal. Fortunately, operating personnel quickly spotted this abnormality and returned the sodium hydroxide to the normal level. The investigation later found the attacker accessed the system via remote access software called TeamViewer, which plant employees had installed and used to check system status and respond to alarms.[49] City officials noted that automated safeguards, such as pH testing, would have triggered an alarm before anyone was harmed, even if the employee had not noticed and stopped the attack.[50] The incident clearly showed, however, that sabotage attacks targeting national critical infrastructure could occur at any moment. For more information on the Oldsmar cyberattack, see chapter 8.

Colonial Pipeline (2021)

Colonial Pipeline, the largest pipeline company in the United States, had to shut down its 5,500-mile pipeline on the east coast for six days due to the ransomware attack by the Russian criminal group called DarkSide.[51] Since the pipeline typically transported more than 110 million gallons of fuel per day, the attack had devastating results: 88 percent of gas stations in Washington, DC, ran out of fuel as did more than 50 percent of gas stations in South Carolina, North Carolina, and Virginia.[52] Although the attack was targeted at IT systems only, the company had to halt its pipeline operation because it could not bill its customers. The fundamental issue

48. "Treasury Sanctions Russian Government Research Institution Connected to the Triton Malware," US Department of Treasury, October 23, 2020, https://home.treasury.gov/news/press-releases/sm1162.

49. CISA, "Alert (AA21-042A): Compromise of U.S. Water Treatment Facility," February 12, 2021, https://us-cert.cisa.gov/ncas/alerts/aa21-042a.

50. Andy Greenberg, "A Hacker Tried to Poison a Florida City's Water Supply, Officials Say," Wired (website), February 8, 2021, https://www.wired.com/story/oldsmar-florida-water-utility-hack.

51. "FBI Statement on Compromise of Colonial Pipeline Networks," Federal Bureau of Investigation (website), May 10, 2021, https://www.fbi.gov/news/pressrel/press-releases/fbi-statement-on-compromise-of-colonial-pipeline-networks.

52. Jonathan Garber, "Colonial Pipeline Fiasco Foreshadows Impact of Biden Energy Policy," Fox Business (website), May 15, 2021, https://www.foxbusiness.com/markets/colonial-pipeline-fiasco-foreshadows-impact-of-biden-energy-policy.

with this incident is that the data necessary for pipeline operations should not be resident on the IT network.[53]

Security Recommendations for ICS

To defend against cyberattacks targeting a critical infrastructure's ICS, organizations need to have good cyber hygiene practices and properly implemented defensive techniques.[54] This section provides an overview of basic cyber hygiene practices and recommended ICS security measures.

Basic Cyber Hygiene Practices

As a fundamental principle of cybersecurity, proper cyber hygiene establishes simple and routine measures to reduce risks from cyber threat actors.[55] In the United Kingdom, a government report in 2015 indicated that 80 percent of cyberattacks could have been prevented if organizations had implemented simple security controls.[56] Although this percentage is not specific to attacks against an organization's ICS, a similar 80-20 rule can be equally applied. Most of the incidents mentioned above were due to inadequate security practices, such as connecting an ICS to the Internet or business network without proper security measures, leaving remote access points open without monitoring, and lack of security controls over removable drives.

There is no clear scope for cyber hygiene. According to a survey on cyber hygiene practices conducted by the European Union Agency for Cybersecurity (ENISA), cyber hygiene generally includes these common practices.[57]

- Identification of hardware and software to determine what to manage.
- Application of secure configuration and hardening for all devices.

53. Joe Weiss, "The Colonial Pipeline Cyberattack—Did IT/OT Convergence Contribute to the Attack," *Control Global* (blog), May 11, 2021, https://www.controlglobal.com/blogs/unfettered/the-colonial-pipeline-cyberattack-did-itot-convergence-contribute-to-the-attack/.

54. "Alert: CrashOverride Malware."

55. ENISA, *Review of Cyber Hygiene Practices* (Athens: ENISA, 2016), 5, https://www.enisa.europa.eu/publications/cyber-hygiene/.

56. Department for Business, Innovation & Skills, "Cyber Security Boost for UK Firms," GOV.UK (website), January 16, 2015, https://www.gov.uk/government/news/cyber-security-boost-for-uk-firms.

57. ENISA, *Review of Cyber Hygiene Practices*, 15.

- Patching systems to keep them current.
- Management of inbound and outbound data.
- Scanning of all incoming e-mails.
- Minimization of the number of administrative accounts.
- Conduct of regular data backup.
- Establishment of an incident response plan.
- Enforcement of security across the supply chain.
- Placement of appropriate security controls in any service agreements.

Similar to these recommended measures in the EU, the US Cybersecurity and Infrastructure Security Agency (CISA) published its Cyber Essential Starter Kit in 2021 to promote basic cyber hygiene practices and a strong culture of cyber readiness. The CISA guide highlights essential steps for an organization to establish cyber readiness in six key areas: management, employees, critical systems, surroundings, data, and an incident response plan.[58]

Essential Cybersecurity Measures Specific to ICS

As Allies and partners consider how to enhance their cybersecurity posture, there are many guidelines, references, and standards that ICS operators and system integrators can refer to for ICS cybersecurity next steps and recommendations. Representing a spectrum of perspectives and best practices employed in various NATO member states, such documents include:

- Canada and the United States: North American Electric Reliability Corporation's Critical Infrastructure Protection Standards[59]
- EU: ENISA's Protecting Industrial Control Systems[60]

58. CISA, *Cyber Essential Starter Kit* (Washington, DC: CISA, 2021), 2, https://www.cisa.gov/sites/default/files/publications/Cyber%20Essentials%20Starter%20Kit_03.12.2021_508_0.pdf.

59. "CIP Standards," North American Electric Reliability Corporation (website), accessed on November 5, 2021, https://www.nerc.com/pa/Stand/Pages/CIPStandards.aspx.

60. ENISA, *Protecting Industrial Control Systems* (Athens: ENISA, 2011), https://www.enisa.europa.eu/publications/protecting-industrial-control-systems.-recommendations-for-europe-and-member-states/.

- France: The National Cybersecurity Agency's (ANSSI) Detailed Measures: Cybersecurity for Industrial Control Systems[61]
- Germany: Federal Office for Information Security's (BSI) ICS Security Compendium[62]
- International: International Electrotechnical Commission 62443 standard series, which currently includes nine standards, technical reports, and technical specifications to secure industrial automation and control systems[63]
- United States:
  - DHS's Catalog of Control Systems Security: Recommendations for Standards Developers[64]
  - DHS's Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies[65]
  - National Institute of Standards and Technology's Guide to Industrial Control Systems Security[66]

As these documents contain vast amounts of information, it is not feasible to examine them more thoroughly in this chapter. Instead, a more helpful framework for Allies and partners seeking to strengthen the security and resilience of ICSs in their critical infrastructure is the *Seven Steps to Effectively Defend ICSs*. After assessing the nearly 300 reported cyber intrusions in 2015, this DHS report identifies seven essential security principles that could have

61. *Agence nationale de la sécurité des systèmes d'information* (ANSSI), *Detailed Measures: Cybersecurity for Industrial Control Systems* (Paris: ANSSI, 2014), https://www.ssi.gouv.fr/uploads/2014/01/industrial_security_WG_detailed_measures.pdf.

62. BSI, *ICS Security Compendium* (Bonn: BSI, 2013), https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/ICS/ICS-Security_compendium.html.

63. "Understanding IEC 62443," International Electrotechnical Commission (website), February 26, 2021, https://www.iec.ch/blog/understanding-iec-62443.

64. DHS, *Catalog of Control Systems Security: Recommendations for Standards Developers* (Washington, DC: DHS, 2011), https://us-cert.cisa.gov/sites/default/files/documents/CatalogofRecommendationsVer7.pdf.

65. DHS, *Recommended Practice.*

66. NIST, *Guide to Industrial Control Systems.*

prevented 98 percent of these incidents.[67] The principles and corresponding security measures outlined in the DHS report are listed below.[68]

- Implement application whitelisting. This step allows only applications and programs predesignated by an administrator to execute, effectively preventing the execution of malware.
- Ensure proper configuration and patch management. Since unpatched systems are more vulnerable to adversaries, this step emphasizes the import and implementation of trusted patches. It includes tracking required patches for each IT asset, obtaining updates from verified sources, validating their authenticity against digital signatures and hash values, testing them on a system equipped with malware detection features, and limiting the connection of external laptops to ICS.
- Reduce attack surface areas. To minimize vulnerabilities, this step seeks to isolate the ICS network from any untrusted networks, lock down all unused ports, disable all unused services, limit external connectivity, use one-way communications for external connectivity if applicable, and employ measures such as restricting a network port or path when bidirectional communications are necessary.
- Build a defendable environment. To limit damages due to breaches of the network, this step calls for segmenting networks into smaller logical enclaves (virtual LANs), restricting host-to-host communications paths, and using a secure means for data transfer from control networks to business networks.
- Manage authentications. Since adversaries seek to gain control of legitimate credentials, this step aims to limit this illegitimate access. Key steps include implementing multifactor authentications when possible, granting users the fewest privileges required to complete duties, enforcing strong password management policies, and not sharing authentication servers between ICS and business networks when centralized authentication is used.

67. DHS, *Seven Steps to Effectively Defend ICSs* (Washington, DC: DHS, 2015), 1–2, https://us-cert.cisa.gov/sites/default/files/documents/Seven%20Steps%20to%20Effectively%20Defend%20Industrial%20Control%20Systems_S508C.pdf.

68. DHS, *Seven Steps to Defend ICSs*, 2–5.

- Implement secure remote access. To counter adversarial attempts to gain unauthorized access to ICSs, this step aims to remove remote access wherever possible. Important actions include limiting any access that remains continuously, implementing read-only access using hardware-type unidirectional network devices, requiring remote access to be time limited and controlled by operating personnel, applying the same remote access paths for vendors and employees, and using two-factor authentication with different types of tokens.
- Monitor and respond. In the modern cyber operating environment, active monitoring is essential. This step recommends monitoring Internet Protocol traffic on ICS boundaries and within the ICS networks, using host-based security solutions to detect malicious software, reviewing login activities to detect stolen credential usage, monitoring changes in access controls, and establishing a sound response plan.

Regarding current threats and vulnerabilities, and the corresponding security measures to mitigate them, various organizations worldwide—including cybersecurity authorities, computer emergency response teams, computer security incident response teams, ICS vendors, and security companies—are continuously issuing advisories, warnings, alerts, and reports. ICS operators and system integrators can stay up to date on evolving cyber threats and appropriate security measures by referencing these documents.

Risk Management for ICS Cybersecurity

The process of risk management is a fundamental task to achieve cybersecurity because it can identify assets that are exposed to risks, assess the level of these risks, implement appropriate measures commensurate with the levels of risks, and continuously monitor and manage the effectiveness of these mitigation steps. When considering risk management practices for IT systems in general—not ICSs in particular—perhaps the most authoritative standard document is *Information Security Risk Management* (ISO/IEC 27005).[69] This document supplements *Information Security Management Systems—Requirements,* the international standard for establishing,

69. *Information Security Risk Management,* ISO/IEC 27005 (Geneva: International Organization for Standardization, 2018), https://www.iso.org/obp/ui/#iso:std:iso-iec:27005:ed-3:v1:en.

implementing, operating, monitoring, and maintaining IT security.[70] As depicted in *Information Security Risk Management*, the general risk management process consists of five essential steps, which are outlined below.

- Context establishment. This steps involves preparation activities, such as setting basic criteria, defining the scope, and establishing a risk management team. The basic criteria include risk evaluation criteria (how to evaluate risks), impact criteria (how to measure impacts), and risk acceptance criteria (thresholds for a desired target level of risk).
- Risk identification. This stage begins with the identification of assets, to include hardware, software, data, information, systems, and process. Then it proceeds with identifying the following information: threats to these assets, existing countermeasures, vulnerabilities that threats can exploit, and potential consequences or damage that could result.
- Risk analysis. This step can be performed in varying degrees of detail. Its methodology can be qualitative—the magnitude and likelihood of an incident are described as low, medium, or high—or quantitative, which uses numerical values rather than descriptions. A combination of likelihood and consequence determines the level of risk for each incident.
- Risk evaluation. This stage helps determine whether risk treatment activities should be carried out for each risk and prioritizes the activities in order of risk level.
- Risk treatment. There are four options available for risk treatment. First, risk modification looks to implement security measures to mitigate risks to an acceptable level by referencing a set of standards and best practices. Next, risk retention accepts risks only when the consequences are negligible or within a range of tolerated outcomes. Third, risk avoidance leads stakeholders to change conditions or cease activities that encounter risks. Finally, risk sharing employs methods like insurance to prepare for residual risks that remain.

Information Security Risk Management recommends organizations perform the risk management process iteratively, starting from an initial high-level assessment to succinctly identify the most critical risks with a broader view.

70. *Information Security Management Systems—Requirements*, ISO/IEC 27001 (Geneva: International Organization for Standardization, 2013), https://www.iso.org/obp/ui/#iso:std:iso-iec:27001:ed-2:v1:en.

Organizations should then perform a detailed assessment that comprehensively analyzes assets, vulnerabilities, threats, and consequences in the second iteration and beyond. Furthermore, organizations should perform risk management regularly, given the evolving nature of the modern security environment. See the thorough explanation of the risk assessment and management process outlined in chapter 13.

Risk Assessment Methodology for ICS

In 2020, the International Electrotechnical Commission published the international standard for ICS risk assessment—*Security Risk Assessment for System Design* (IEC 62443-3-2)—and adopted it as part of the broader *Security for Industrial Automation and Control Systems series*.[71] A key concept in *Security Risk Assessment for System Design* is the consideration of ICS zones and conduits. A zone is a collection of logical and physical assets posing the same characteristics from the perspective of security requirements, criticality, and logical and physical relationships. A conduit is a logical grouping of communications channels that have the same security requirements, and each conduit represents the connection between two or more zones.

Another distinguishing aspect of IEC 62443-3-2 is that it utilizes the concept of security level (SL)—a measure of confidence that the ICS is free from vulnerabilities and is functioning as intended—to assist organizations in identifying required security measures. Derived from the international standard *System Security Requirements and Security Levels* (IEC 62443-3-3), a standard practice is to assign a label to each security measure ranging from SL1 (basic security) to SL4 (most sophisticated security).[72] After assigning these labels, organizations then use them to identify recommended security measures commensurate with their target level of protection. For example, as for the security requirements related to "system log storage capacity," IEC 62443-3-3 suggests that using a storage with sufficient capacity would be just sufficient for SL1 and that a warning function against low disk space should be added to achieve SL2 or above.

Similar to the iterative approach used in *Information Security Risk Management,* the ICS risk assessment process outlined in IEC 62443-3-2 is also divided into two levels, namely, initial risk assessment and detailed

71. *Security for Industrial Automation and Control Systems—Part 3-2: Security Risk Assessment for System Design*, IEC 62443-3-2 (Geneva: International Electrotechnical Commission, 2020), https://webstore.iec.ch/publication/30727.

72. *Industrial Communication Networks—Network and System Security—Part 3-3: System Security Requirements and Security Levels*, IEC 62443-3-3 (Geneva: International Electrotechnical Commission, 2013), https://webstore.iec.ch/publication/7033.

risk assessment. The process for ICS risk assessment consists of seven steps, which are described below and illustrated in figure 14-2.

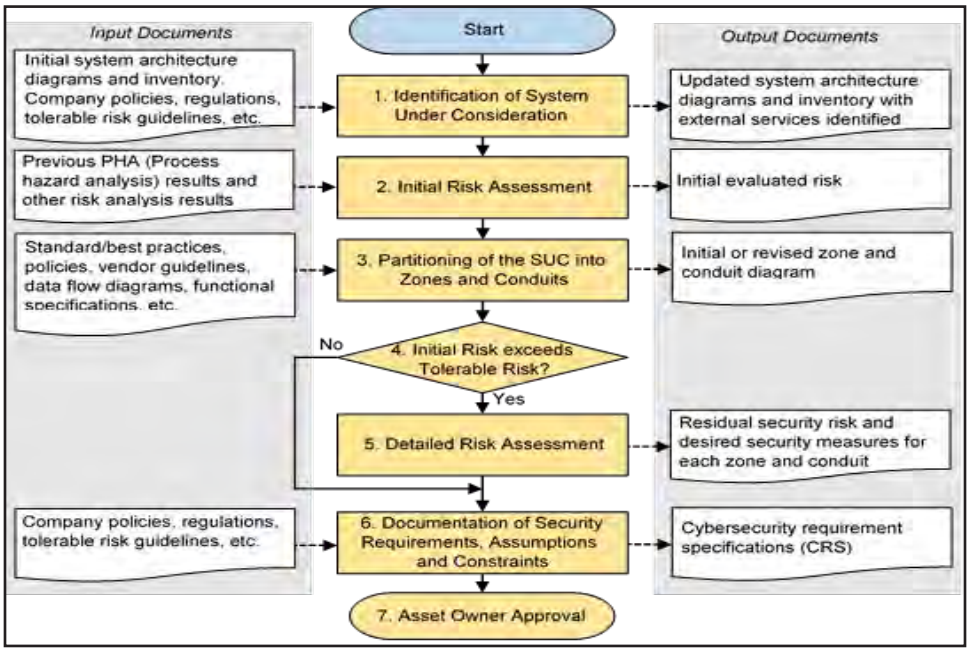


Figure 14-2. Workflow diagram for ICS risk management
(Diagram by the International Electrotechnical Commission)

- Identification of system under consideration (SUC). Step 1 identifies the SUC, including identification of the ICS boundary, access points, and all ICS assets.
- Initial risk assessment. Step 2 identifies the worst-case scenarios by assuming the likelihood of occurrence to be 100 percent certain. The purpose of the initial assessment is to identify and prioritize the areas for detailed assessments.
- Partitioning of the SUC into zones and conduits. Step 3 includes a grouping of ICS assets based on the initial assessment results so that assets with the same characteristics are grouped into the same zones. Organizations are recommended to group unordinary devices (such as wireless devices and devices connected to external networks) into separate zones because they require special care.

- Risk comparison. In Step 4, organizations determine if an additional detailed risk assessment is required for the SUC (or part of it) by comparing the initial assessed risk to the level of risk the organization can tolerate. If the assessed risk exceeds the tolerable risk, then the organization should perform a detailed risk assessment.
- Detailed risk assessment. Step 5 builds on the previous steps and goes into greater examination of the system, using a series of micro-steps. Here, organizations (1) identify all threats that could affect the assets within the zone or conduit, (2) identify areas in which assets are vulnerable to these threats, (3) develop a worst-case estimate of potential impacts, (4) estimate the likelihood of such incidents occurring, (5) assess the level of risk for each threat, (6) compare the assessed risk to the tolerable risk to determine whether to accept, transfer, or mitigate the risk, (7) assess residual risks that remain after applying mitigation measures, and (8) identify additional measures when the residual risks exceed the tolerable risks.
- Documentation of security requirements, assumptions, and constraints. Step 6 is about documenting all the findings from previous steps. The cybersecurity requirements specification contains the description of mandatory security measures as well as details of the SUC, zones and conduits, threat environments, organizational policies, and tolerable risks.
- Asset owner approval. At the final step of each iteration of risk assessment, asset owners in charge of the safety and reliability of control processes review and approve the result.

Detailed Risk Assessment Approach

Since it provides an in-depth understanding of the nature of risks, a detailed approach to risk assessment is at the heart of managing risks to ICS and securing them more effectively. The risk assessment process requires an organization to estimate the likelihood of a threat and impacts of potential incidents for every pair of assets and threats. This process can be tedious and consume time and resources because it requires tremendous effort to have meaningful and valid results.

Qualitative or descriptive measures such as *high*, *medium*, and *low* can be used in estimation; however, they still require a detailed guideline to reduce subjective and ambiguous judgments as much as possible. Moreover, the impact has multiple attributes, requiring in-depth review from various perspectives, such as an outage of service, loss of process accuracy, and the impacts on health, safety, and environment. Once the organization estimates the maximum potential magnitude of impact and its likelihood for every asset-threat pairing, the organization can then determine the level of risk.

Qualitative analysis, on the other hand, uses simple mapping logics to determine the risk level. For instance, the risk is high if the likelihood and impact are both assessed as high. These logics are generally expressed in a matrix, called the *risk matrix*, with rows representing qualitative values for likelihoods and columns representing qualitative values of impacts. Quantitative analysis uses numerical metrics (such as annual loss expectancy), which is the monetary loss amount multiplied by the probability of occurrence. The measurement and assessment of risks serve as a basis of deciding which risks to prioritize in order of importance.

Scenario-based Approach for Security Baseline

For organizations with no experience or expertise in detailed risk assessment, the scenario-based risk mitigation approach may be helpful as a starting point toward developing more robust and effective ICS cybersecurity. This approach considers past incident cases or potential scenarios to identify required security measures to prevent such incidents from occurring. It can also be used to assess the effectiveness of the current security posture with relatively less time and resources.

Under the scenario-based approach, an organization first needs to identify assets, zones, and conduits within the ICS, and then build a catalog of threat scenarios applicable to them. To build a quality catalog of threat categories, organizations can compile incident reports and security warnings or advisories from various sources. Then, the organization should identify required security measures by referencing best practices and standards or by brainstorming with relevant stakeholders and experts. The next step is to evaluate the feasibility of the identified security measures. When any specific security measure cannot be implemented due to budget or technical restrictions, the organization should seek alternative or compensating controls (such as adding manual control procedures or physical controls).

The key advantage of the scenario-based approach is that no additional analysis skills are required for risk mitigation, so organizations can complete this type of assessment more quickly than a detailed risk assessment. The main disadvantage, however, is that some important risks could be overlooked, especially those risk scenarios that have not occurred elsewhere and thus are not considered as being in the realm of possible.[73] Another disadvantage is that there is little justification for chosen security measures from the viewpoint of cost-effectiveness because this approach does not consider the impact and likelihood of incidents. A robust catalog of threat scenarios could reduce these shortcomings to some extent.

When building a catalog of threat scenarios, Allies and partners may find the MITRE Corporation's Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) for Industrial Control System (ICS) quite helpful as a tool and guide.[74] MITRE started building the ATT&CK for IT systems in 2013 and it is now widely accepted as a framework for documenting and analyzing tactics and techniques used by cyberattackers. MITRE's ATT&CK for ICS, launched in 2020, contains details of 78 attack techniques that threat actors employed in the wild along with corresponding mitigation measures organizations can take to enhance their cybersecurity posture. Another valuable source of information is the Top 10 Threats and Countermeasures for ICS, which the German BSI began publishing in 2014 to highlight the most severe but common cyber threats and outline appropriate security measures for organizations to adopt.[75]

Defending against Cyberattacks: Looking to the Future

The steps critical infrastructure owners and operators take to manage security risks and threats in their respective operational environments are vital to achieving cybersecurity. Their governments should also play a proactive role to build resilience and prepare for potential cyberattacks at both the national and international levels. The following section discusses the important efforts governments should undertake.

73. *Industrial Communication Networks—Network and System Security—Part 2-1: Establishing an Industrial Automation and Control System Security*, IEC 62443-2-1 (Geneva: International Electrotechnical Commission, 2010), 48, https://webstore.iec.ch/publication/7030.

74. "Techniques," MITRE Corporation (website), January 2, 2020, https://collaborate.mitre.org/attackics/index.php/All_Techniques.

75. BSI, *Industrial Control System Security Top 10 Threats and Countermeasures* (Bonn: BSI, 2019), https://www.allianz-fuer-cybersicherheit.de/SharedDocs/Downloads/Webs/ACS/DE/BSI-CS/BSI-CS_005E.pdf.

National-level Efforts for CISR

To varying degrees, national governments conduct cybersecurity governance activities through central ministries or authorities, and develop and update their respective cybersecurity strategies that stipulate necessary measures to protect critical national infrastructure. In 2016, ENISA published a list of good practices through a detailed analysis of various governance activities across 15 EU member states. Some of the key practices that the report recommends EU member states adopt are listed below.[76]

- Partnerships with private stakeholders. As private companies manage many critical infrastructure systems and assets, it is essential to have a strong partnership between the government and the private sector in an institutional form, such as a national critical infrastructure protection committee or advisory meeting. See chapter 11 for its recommendations for public-private partnerships.
- Information-sharing scheme. Cyber threat information should be disseminated to all relevant government agencies and private critical infrastructure operators through preestablished information-sharing schemes. These established procedures allow relevant stakeholders to obtain up-to-date information promptly and take appropriate security measures.
- Develop the community of computer security incident response teams. Establishing the institutional foundation for cooperation among public and private response teams can lead to mutual benefits, such as increased knowledge and more efficient allocation of resources.
- Risk assessment. The government should guide and support private operators to identify risks and implement security measures as requested.
- Cyber crisis management. Cyber crisis management should include the definition of roles and responsibilities, and decision-making procedures between relevant stakeholders.

76. ENISA, *Stocktaking, Analysis and Recommendations on the Protection of CIIs* (Athens: ENISA, 2016), 16–19, https://www.enisa.europa.eu/publications/stocktaking-analysis-and-recommendations-on-the-protection-of-ciis/.

- Comprehensive legal framework. Countries should have laws and regulations pertaining to securing critical infrastructure that stipulate the mandatory requirements for implementing essential security measures and notification of cyber incidents.

All of the steps listed above are important, but the most vital practice is information sharing. Private operators generally do not want their incidents to be disclosed to the public, while national and military intelligence agencies typically are reluctant to share their confidential information with the private sector. To overcome this problem, it is necessary to build trust that the shared information will never be leaked to other parties. The government should establish a formal information-sharing policy, including a sanitization process to remove sensitive content when disseminating information from a specific operator to other operators. Additionally, signing a nondisclosure agreement between parties can also build trust. For an in-depth discussion on helpful information- and intelligence-sharing practices, see chapter 11.

For effective and efficient dissemination of information, the government should use IT-based communications means. Depending on the size of the country, there may be thousands of critical infrastructure facilities, making the timely dissemination of threat information to all owners and operators almost impossible with manual handling procedures. In most situations, information technologies provide a more efficient and timely venue for multidirectional information sharing between the government and all relevant stakeholders. There are two examples of such programs used in the United States: the Homeland Security Information Network (HSIN) and the Automated Indicator Sharing (AIS), operated by the DHS and the CISA, respectively. The HSIN is an information portal for trusted information sharing between federal, state, local, international, and private-sector partners.[77] In contrast to the HSIN, the AIS is a real-time automated dissemination mechanism that sends machine-readable cyber threat indicators of compromise—artifacts observed on a network or operating system that indicate a cyber intrusion—to the participants of the AIS community.[78] Examples of these indicators include Internet Protocol addresses, domain names of C2 servers, and hash values of malware.

Beyond the recommendations in the ENISA report, two additional best practices are the use of cyber exercises and supply-chain security. Since critical infrastructure systems and sectors are highly interrelated,

77. "Homeland Security Information Network (HSIN)," DHS (website), December 3, 2021, https://www.dhs.gov/homeland-security-information-network-hsin.

78. "Automated Indicator Sharing," CISA (website), accessed on October 23, 2021, https://www.cisa.gov/ais.

an attack on a particular facility can affect other infrastructures rather than simply being confined to the initial target of the attack. In particular, attacks against the lifeline sectors (such as electricity and telecommunications) may affect all other sectors. To prepare for national-level cyber crises, the government should host exercises regularly with all relevant stakeholders. These exercises should include the procedures of decision making and communications across all government areas as well as the procedures for individual operators to respond to cyberattacks and report them to the government. *Cyber Storm*, which focuses on cyberattack crisis management, is the largest cyber exercise in the United States.[79] Similarly, *Cyber Europe i*s a large-scale cyber exercise that tests procedures, communications, and decision making at the EU level.[80]

Supply-chain security is a relatively new area of concern. The supply chain of hardware and software used for critical infrastructure should be protected against intentional and accidental modification that could be incurred during entire life cycles of products, including development, delivery, and maintenance. Malicious interference by a nation-state in cooperation with manufacturers located in its territory—by implanting a backdoor within IT/OT components, for example—is incredibly difficult to discover. Moreover, criminal or terrorist groups can also cause harm to IT/OT components by infiltrating manufacturers' development environments to modify source codes. Therefore, the government should establish a framework to screen the trustworthiness of manufacturers and ensure the security of products for their entire life cycles.

International-level Efforts for CISR

International cooperation is also essential to protect critical infrastructure because of the borderless nature of cyberspace. It is almost impossible for a single country to thoroughly analyze cross-border attacks and block further ones because attacks generally take place over multiple stages across several countries. Moreover, one country may possess intelligence that another country does not have. A complete analysis, investigation, and attribution of an attack thus require close international cooperation. Ideally, all government agencies involved in securing critical infrastructure (such as the national cybersecurity authority, national and military intelligence agencies, cyber commands, law enforcement agencies, and computer security

79. "Cyber Storm: Securing Cyber Space," CISA (website), accessed November 3, 2021, https://www.cisa.gov/cyber-storm-securing-cyber-space.

80. "Cyber Europe," ENISA (website), accessed on November 3, 2021, https://www.enisa.europa.eu/topics/cyber-exercises/cyber-europe-programme.

incident response teams) should have close international cooperation channels with relevant counterparts in foreign countries. Multilateral treaties or agreements—like the Convention on Cybercrime of the Council of Europe, also known as the Budapest Convention—can play a crucial role since all members would be obliged to cooperate without having to make separate bilateral agreements with each other.[81] The Budapest Convention, currently signed by 66 countries, is the international treaty on cybercrimes to obtain a series of powers and procedures required for law enforcement. Article 23 of the convention stipulates that international cooperation is to be provided among participants to the widest extent possible.

Likewise, participating in international malware information-sharing platforms (MISP), such as the MISP sponsored by NATO and the EU, will provide the participating countries with up-to-date global threat information and relevant indicators of compromise on a real-time basis.[82] MISP is an open-source information-sharing platform developed by a team of cybersecurity experts from the Computer Incident Response Center in Luxembourg, the Belgian Ministry of Defense, and NATO. MISP can share, store, and correlate indicators of compromise, threat intelligence, vulnerability information, and even counterterrorism information.[83] Allies and partners may also consider participating in HSIN and AIS, as access can be granted to non-US entities under certain conditions.

Areas for international cooperation are not limited to exchanging threat information, sharing intelligence, and supporting investigations. Instead, it should include exchanges of various cybersecurity know-how and best practices, such as lessons learned from certain types of cyber incidents, detailed information on technical cybersecurity measures, policies for securing supply chains against cyber threats, and tools for assessing an organization's cybersecurity level. As countries exchange such information and provide technical support and consultation to one another, if requested, their cooperation will help build common capabilities to achieve cyber security, defense, and resilience at sufficient levels to secure critical national infrastructure. International cooperation is of paramount importance in general, but especially for EU member states, because many

81. "Details of Treaty No. 185," Council of Europe (website), accessed November 3, 2021, https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185.

82. "Who Is behind the MISP Project?," MISP Threat Sharing (website), accessed on November 5, 2021, https://www.misp-project.org/who/.

83. "What Is Malware Information Sharing Platform (MISP)?," Cyware (website), September 17, 2020, https://cyware.com/educational-guides/cyber-threat-intelligence/what-is-malware-information-sharing-platform-misp-b28e.

European critical infrastructure sectors and systems are interconnected. The European power grid as well as oil and gas pipelines are two key examples of this connectivity.[84] An incident in one country may affect other countries, potentially leading to a cascade effect. See chapter 12 for more detail on the nature of dependencies and interdependencies among critical infrastructure sectors.[85]

Conclusion

This chapter provided a brief overview of the characteristics of ICSs, major cyber incidents against ICSs, essential cybersecurity measures, and risk management methodologies. Cyber incidents against critical infrastructure continue to occur due to inadequate security management practices, system misconfigurations, and human errors. Since critical infrastructure plays an important role in social well-being and national security, operators should maintain a sense of mission to cybersecurity, keep vigilant against cyberattacks and incidents, and make continuous efforts to strengthen the systems.

Governments should also make tremendous efforts to protect their critical infrastructure by establishing mandatory security requirements for critical infrastructure, ensuring owners and operators comply with these requirements, and providing security advice as needed. In addition, governments should be transparent about security matters and promptly share threat information with the critical infrastructure operators.

Government organizations, security companies, and manufacturers have different capabilities and specialties. It is, therefore, necessary to create an institutional cooperation mechanism (such as a public-private critical infrastructure security council and a joint cyber response team) so stakeholders' unique capabilities can be integrated at the national level. Each country should also build trust with international partners and actively share information and intelligence. This cooperation will allow like-minded countries not only to detect, prevent, and investigate attacks in a timely manner, but also to build a framework for international collaboration in which they can work together to improve cybersecurity and resilience,

84. "ENTSOE-E Transmission System Map," ENTSO-E (website), January 1, 2019, https://www.entsoe.eu/data/map/; and "Europe Pipelines Map," Theodora (website), March 31, 2017, https://www.theodora.com/pipelines/europe_oil_gas_and_products_pipelines.html.

85. ENISA, *Communication Network Dependencies*, 23–24.

determine attribution for cyberattacks, and take harmonized actions against threat actors who perpetrate them.

– 15 –

Crisis Management and Response

Malcolm Baker

The North Atlantic Treaty Organization describes *crisis management* as the "coordinated actions taken to defuse crises, prevent their escalation into armed conflict and contain hostilities if they should result."[1] As such, crisis management is an essential component of the Alliance's Strengthened Resilience Commitment announced in June 2021 as part of the NATO 2030 initiative. Integral to this commitment is the philosophy of Article 3—Allies' individual and collective capacity to resist armed attack—and NATO's ability to fulfill its three core tasks of collective defense, cooperative security, and crisis management.[2] This commitment reemphasizes that resilience within NATO member states and partner countries is both a "national responsibility and a collective commitment."[3] Specific to critical infrastructure, the official announcement of NATO 2030 also references increasing efforts to "ensure the resilience of our critical infrastructure (on land, at sea, in space and cyberspace) and key industries, including by protecting them from harmful economic activities."[4]

1. North Atlantic Treaty Organization (NATO) Standardization Office, *NATO Glossary of Terms and Definitions,* AAP-06, Edition 2020 (Brussels: NATO, 2020), 36.

2. "Crisis Management," NATO (website), October 8, 2020, https://www.nato.int/cps/en/natohq/topics_49192.htm.

3. "Strengthened Resilience Commitment," NATO (website), June 15, 2021, https://www.nato.int/cps/en/natohq/official_texts_185340.htm?selectedLocale=e.

4. "Strengthened Resilience Commitment."

Within the overall construct of resilience and NATO's comprehensive approach to its core missions, however, is the Alliance's current philosophy of crisis management keeping up with mainstream developments in contemporary crisis management and thought leadership? More importantly, within the construct of critical infrastructure security and resilience (CISR) efforts, is the NATO 2030 approach to crisis management still fit for purpose or could it be improved to meet the future challenges of interconnected critical infrastructure systems?

In response to these questions, this chapter will examine the broad issue of crisis management within the context of CISR programs at the Alliance's organizational level and among its Allies and partners. For example, how does crisis management fit within the broader concept of resilience and therefore CISR? The chapter will also explore how effective CISR measures can be improved further by developing and implementing robust crisis management structures and processes. Therefore, it is important when considering CISR and crisis management to focus on the role that critical infrastructure plays within NATO, its member states and partner nations, and the potential impacts or consequences to NATO operations that can arise if critical infrastructure services are hindered or otherwise unavailable. Finally, the chapter will review new developments and emerging themes in resilience and crisis management, and offer suggestions for how NATO could better align its activities in this discipline to support NATO 2030.

Critical Infrastructure

As chapter 1 outlined in the beginning of this book, NATO Allied Command Operations defines critical infrastructure as "a nation's infrastructure assets, facilities, systems, networks and processes that support the military, economic, political and/or social life on which a nation and/or NATO depends."[5] This definition presents three different subcategories of critical infrastructure—critical national infrastructure, mission-vital infrastructure, and key infrastructure—to communicate the importance of these facilities or services to national security and/or Allied operations.

By its definition, critical infrastructure implies the essential supply of services to a nation-state and its economy, communities, and citizens. It follows, therefore, that any disruptive event that adversely affects critical infrastructure will result in corresponding negative impacts on the continued supply of these vital services. Many modern critical infrastructure facilities

5. Allied Command Operations (ACO), *Infrastructure Assessment*, ACO Directive 084-002 (Mons, BE: ACO, 2019), 4.

are highly dependent and interdependent on other critical infrastructure sectors, and they rely upon power, telecommunications, and other services—such as emergency services, law enforcement and potentially local, regional, and state utility services—to function properly and efficiently. See the in-depth discussion of dependencies and interdependencies in chapter 12.

Critical infrastructures are often complex, and over time they have become more resilient to threats, hazards, and risks in general. State and national programs have made progress in enhancing security of critical national infrastructure assets and increasing their resilience to incidents, emergencies, and crises. Each sector, however, tends to build its resilience in isolation rather than holistically—a practice driven in part by those sectors subject to state regulation.

Why Is Crisis Management and Response Important?

Since it provides services that impact modern life and national security, critical infrastructure is economically important to NATO member states and partner nations and to the citizens who rely on these services. Critical infrastructure may be owned either by a state or by a corporate entity in the private sector. In many cases, the state regulates or governs critical infrastructure in terms of how it operates, what it produces, or what prices it charges for the good and services it delivers. Therefore, chief executives, board members, and company directors of a critical infrastructure facility may answer to outside entities such as the state, government departments, regulators, or investors.

Depending on the nature of the critical infrastructure or the sector in which it operates, there may also be international treaties or other multilateral agreements or conventions that govern how the organization or facility should maintain services, safety, and security obligations. Such treaties and conventions also place similar obligations on critical infrastructure owners, managers, and operators, beyond their traditional roles and responsibilities. Directors and board members may well have fiduciary duties to manage organizations effectively and efficiently, including crisis management and response plans when things go wrong.

Incidents, Emergencies, and Crises: What Is the Difference?

The terms *incident*, *emergency*, and *crisis* are often used erroneously as interchangeable or synonymous words to describe an event that occurs or a scenario that unfolds. Although different, each of these terms is used to explain negative or unwelcome consequences. Incidents, emergencies, and crises can bring about not only negative consequences, but they can also sometimes positive outcomes or opportunities for organizations.

Academics have long written about crises and crisis management and developed definitions for what constitutes a crisis. NATO, however, does not have an agreed or published definition of a crisis. If NATO and its member states and partners are to develop crisis management structures and processes, then it is vital to share a common understanding of the term crisis. Specifically, what types of scenarios are the Alliance, member states, and partner nations seeking to manage, and how can these events—whether they are incidents, emergencies, or crises—be managed or resolved and their consequences be mitigated or minimized? Is there a one-size-fits-all solution that could manage all events?

The academic literature offers multiple definitions of what makes a crisis. Based on his extensive literature review, Patrick Lagadec describes the anatomy of a crisis as a combination of uncertainty, unknowns, the unthinkable, the unimaginable, and the unforeseeable.[6] Steven Fink defines a crisis as "an unstable time or state of affairs in which a decisive change is impending—either one with the distinct possibility of a highly undesirable outcome or one with a distinct possibility of a highly desirable and extremely positive outcome."[7] Another definition describes a crisis as "a damaging event or series of events that display emergent properties which exceed an organization's abilities to cope with the task demands that it generates and has implications that can effect a considerable proportion of the organization as well as other bodies." Denis Smith and Dominic Elliott explain further how crises can manifest or are triggered by internal or external incidents that expose an inherent latent vulnerability embedded with the organization.[8] These definitions are only three among a myriad of others offered in the

6. For additional information, see Patrick Lagadec, *Preventing Chaos in a Crisis: Strategies for Prevention, Control and Damage Limitation* (London: McGraw-Hill, 1993).

7. Steven Fink, *Crisis Management: Planning for the Inevitable* (Lincoln, NE: iUniverse Inc., 2002), 15.

8. Denis Smith and Dominic Elliott, *Key Readings in Crisis Management: Systems and Structures for Prevention and Recovery* (New York: Routledge, Taylor & Francis Group, 2006).

academic literature, but Lagadec suggests personnel responsible for responding to, and managing, a crisis need a more practical, less theoretical description.[9]

Therefore, in search of a more practical and real-world definition that suits the needs of readers, table 15-1 offers the following definitions based on published crisis management guidance and best practices in the United Kingdom.[10]

Table 15-1. Key definitions in UK crisis management

Incident	An adverse situation that might cause disruption, loss or emergency, but which does not meet the organization's threshold, or definition of, a crisis.
Emergency	An incident that requires an immediate response to minimize loss of life or serious injury/harm; or serious damage to property.
Crisis	An inherently complex, abnormal and unstable situation that due to its scale, duration and impact threatens the organization's strategic objectives, operations, reputation or viability, or has strategic implications for the organization.
Crisis management	The developed capability of an organization(s) to prepare for, anticipate, respond to, and recover from crises.

These definitions are more likely to be understood by owners and operators of critical infrastructure. Additionally, the definition of crisis management here is more appropriate and relevant to critical infrastructure than the NATO definition mentioned at the beginning of the chapter. The commitment to strengthen resilience, which undergirds NATO 2030, envisions a broad approach working across the "whole of government, with the private and non-governmental sectors, with programmes and centres of expertise on resilience established by Allies, and with our societies and populations, to strengthen the resilience of our nations and societies."[11] Given the Alliance's pursuit of the NATO 2030 initiative, understanding crisis management exclusively in terms of armed conflict and other hostilities may no longer be appropriate or optimal, especially in when considering the various physical, cyber, and hybrid threats outlined earlier in chapters 2–4. Therefore, NATO 2030 presents an opportunity to renew the Alliance's

9. For additional information, see Lagadec, *Preventing Chaos in a Crisis.*

10. British Standards Institution (BSI), *Crisis Management—Guidance and Good Practice: BS 11200: 2014* (London: BSI Standards Limited, 2014), 2.

11. "Strengthened Resilience Commitment."

efforts to develop greater resilience and update how it views and exercises civil-military cooperation and crisis management.

The definitions in table 1, based on a NATO member state's government-sponsored standards, may prove helpful and useful in supporting NATO 2030. Developed as a result of learning from the experiences of managing complex, interconnected, and interdependent crises over a long period of time, they represent a modern and refreshed approach to resilience and crisis management.

It is also beneficial to examine the nature of crises so that when critical infrastructure stakeholders and decisionmakers face such events, they understand the typical characteristics of a crisis and are better postured to respond effectively. Table 15-2 highlights the unique nature of crises by contrasting them with incidents or emergencies on the basis of six characteristics.[12]

Table 15-2. Differences between crises and incidents or emergencies

Characteristics	Incidents/Emergencies	Crises
Predictability	▪ Generally foreseeable and managed with predetermined response or contingency plans ▪ The timing, extent, type of incident, and its impact is variable and unpredictable in terms of detail	▪ Often rare, unpredictable, and inherently complex and unstable ▪ May escalate from the mismanagement of events ▪ Pose strategic challenges that threaten an organization's survival ▪ Do not respond to emergency plans and predetermined responses
Onset	▪ Can occur suddenly with little or no notice ▪ May emerge over time due to a gradual failure or latent defect in systems and processes ▪ Sometimes indicators and warning of an issue can be monitored and recognized as a potential trigger or actual problem	▪ Can occur suddenly with little or no notice or warning ▪ Can form a rising tide event linked to an incident that escalated and now presents organizational threats and strategic challenges ▪ May manifest due to latent unresolved issues or systemic faults in the organization ▪ Present unparalleled reputational challenges to the organization

12. BSI, *Crisis Management*, 4.

Table 15-2 (continued). Differences between crises and incidents or emergencies

Characteristics	Incidents/Emergencies	Crises
Urgency and pressure	▪ Duration of the event and requisite response are usually short-lived ▪ Timely resolution prevents an incident from exposing longer-term or malign impacts of the organization	▪ Require greater sense of organizational urgency and strategic attention ▪ May manifest over weeks, months, or longer ▪ Organization must work for a longer period to mitigate and minimize impacts
Complexity: Scale, duration, and impacts	▪ Often occur with relative frequency so that they are understood and predictable ▪ Organizations often have predetermined plans and response mechanisms that can be implemented ▪ May have wider impacts on occasion	▪ Can transcend organizational and territorial boundaries ▪ Have potential and propensity to disrupt multiple organizations or different sectors (government, private sector, or communities) ▪ Often associated with uncertainty and little, imprecise, or ambiguous information ▪ True extent and impact may not be known immediately
Reputation: Media scrutiny and public outcry	▪ Mostly positive and supportive when dealt with robustly and professionally and resolved speedily ▪ Can be negative even if they are resolved successfully ▪ Adverse coverage can escalate them into an organizational crisis	▪ Create intense, sustained public attention and media interest ▪ Continued, negative coverage has the potential to harm an organization's reputation ▪ Social media and "citizen journalism" may increase inaccurate reporting ▪ Post-event inquiries, hearings, or trials may prolong media interest
Resolution: Manageability with existing plans and procedures	▪ Often resolved using predetermined plans, procedures, and response arrangements ▪ May include activities to preempt an incident as well as mitigation, and recovery ▪ In most cases adequate resources are available	▪ Seldom resolved using conventional, predetermined plans and procedures due to their inherent complexity, instability, infrequency, and unpredictability ▪ Require flexible and innovative strategic command and control over sustained period of time ▪ May require more resources and over a longer period ▪ Response should achieve strategic objectives of all organizations involved

Based on the characteristics outlined in table 2, it is clear crises are more complex and challenging in terms of their nature and the steps required to respond to them effectively. Figure 15-1 combines parts of different crisis management models to provide a pictorial representation of the anatomy of a crisis and the subsequent response phases.

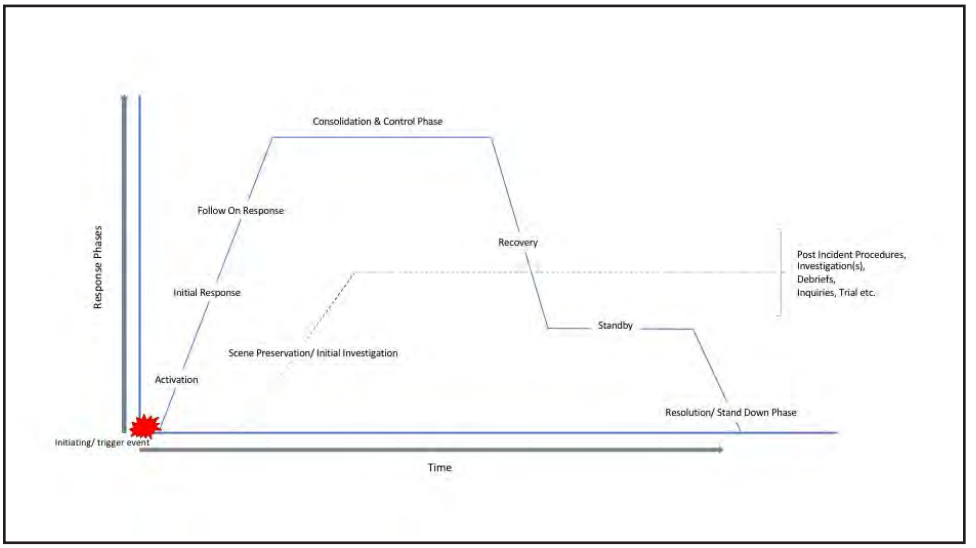


Figure 15-1. Anatomy of a crisis

As figure 15-1 illustrates, the initiating event or trigger starts the response process, but the event could be a slow burn or incremental event rather than one with a sudden impact. The diagram represents a general depiction to illustrate the key phases of a crisis. Achieving consolidation and control over a crisis could take time—a point made all too clear during the outbreak of the COVID-19 pandemic—and would be prolonged further by the public inquiries and hearings that would likely follow.

With this foundation established for understanding the nature of a crisis, it is now appropriate to consider the elements of responding to and managing a crisis. In a helpful framework, Fink identifies the four key stages in crisis management, explained in further detail below.[13]

- Prodromal crisis stage. This is the warning stage, sometimes referred to as the pre-crisis stage. There is not always a pre-crisis stage, as there are times when a crisis arrives without any warning. If there is a warning, however, and

13. Fink, *Crisis Management*, 26.

it is missed, then the next stage is possibly the first indication that an organization is in a crisis. In figure 15-1, this stage roughly equates to the events preceding the initiating event and the trigger event itself.

- Acute crisis stage. This stage is sometimes described as the "point of no return" because if an organization has not reacted during the early warning stage, then it has lost that opportunity. The organization may be able to control some aspects of the crisis, however, by implementing mitigation or damage limitation measures. The acute stage is synonymous with the activation stage, initial and follow-on response phases, and the consolidation phase in figure 15-1.
- Chronic crisis stage. This is the phase in which the investigations, inquiries, and post-event reviews take place, often leading to either recriminations and assigning blame or to acknowledging successes and opportunities.
- Crisis resolution stage. This stage generally reflects the resolution of the crisis and, if successful, it should be as close to the prodromal phase in terms of time. This stage signifies the organization has recovered from the event.

Therefore, if a crisis cannot be managed by predetermined plans and procedures or does not lend itself to successful resolution by applying such measures, then it demonstrates the need for a different crisis management construct. Indeed, a crisis management capability is required. Figure 15-1 as a simplified model is consistent with Fink's framework and answers Lagadec's call for a less theoretical description of crisis management. The key elements of effective crisis management are early warning, an effective strategy, good communication, leadership, and swift decision making.

Developing Crisis Management Capability

During the past decade, many countries have adopted a framework similar to the UK concept of integrated emergency management to develop their respective national resilience programs. Based on potential threats, hazards, and risks, this concept embraces six stages: (1) anticipation, (2) assessment, (3) prevention, (4) preparation, (5) response, and (6) recovery

management.[14] In many ways, this doctrine and approach can be applied to building a crisis management capability, though the capability outputs are different. Rather than developing incident and emergency response plans as the products, applying this process in a similar manner assists in developing a crisis management framework. This framework, in turn, helps establish a management structure and process to manage crises effectively. NATO, as an institution, and the armed forces of its member states, in general, are well versed in the concept and practice of developing capability and capacity. Capability development in crisis management can follow established, recognized good practice in developing capability requirements, identifying capability gaps, and determining the key steps required to build capability and capacity.

Figure 15-2 provides NATO, its member state governments, and private-sector entities (including owners and operators of critical infrastructure) with a proven crisis management framework.[15] Based on the previous examination of the nature of a crisis in table 15-2, this framework does not include the prevention stage found in the six-step concept of integrated emergency management because crises can seldom be prevented. The framework provides an iterative model that suggests continuous improvement to the extent that organizations should always learn from experience. The remainder of this section will discuss each phase or stage of the framework.

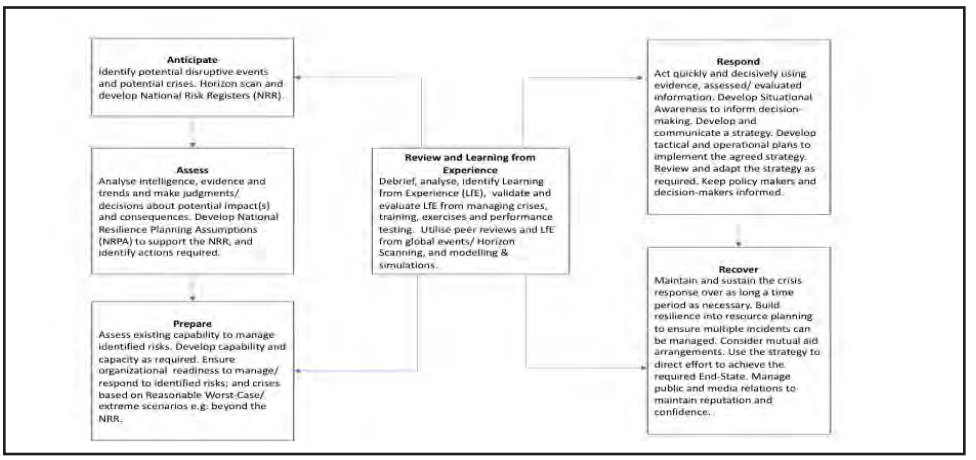


Figure 15-2. A general framework for crisis management
(Diagram by the British Standards Institution)

14. *Civil Contingencies Act Enhancement Programme: Emergency Preparedness* (London: Cabinet Office, 2012), 16–18, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/61024/Chapter-1-Introduction_amends_16042012.pdf.

15. BSI, *Crisis Management*, 9.

Anticipate and Assess

These first two stages in the crisis management framework require an organization (such as NATO, its member states, and partner countries, or critical infrastructure owners and operators) to acknowledge and understand the causal relationships between its strategic objectives, risk management, and identified challenges to the organization. In effect, the organization should have a well-defined and developed risk register. It is likely most Allies or partners have already established a register similar to a National Risk Register (NRR), including a list of identified and assessed risks to the state. The failure to develop a reliable and meaningful appreciation of risk means the organization will be inevitably prone to failure if a crisis occurs.

Typically, an all-hazards, all-threats approach would be appropriate and allow a government to consider the full range of potentially disruptive events it could face. See the discussion of an all-hazards, all-threats approach in chapter 2. This process of anticipation includes identifying risks, assessing likelihood or probability of occurrence, understanding the immediate impact, and extrapolating the impact to understand the longer-term consequences. As part of an organization's risk assessment methodology, there should be a well-developed, rigorous, and demonstrable horizon-scanning process, together with active early warning, that ensures the organization recognizes an identified risk at an early stage.

Horizon scanning provides Allies and partners with a useful tool to assist in the development or updating of a National Risk Assessment (NRA). Global events and crises may trigger or initiate the process of an organization asking itself to what extent it is prepared if such a crisis occurred. To develop NRRs and the NRA, governments typically build multidisciplinary teams, assisted by relevant subject-matter experts according to the topic. In the case of malicious threats, it would be prudent to include national security, intelligence, and law enforcement agencies in the process. Risk registers are built from the bottom up, effectively incorporating the risks organizations and departments face at different levels of government. States then compile and assess the risk registers at a national level to form the NRR and identify the top-tier risks that should be included as the foundation for the NRA. The NRRs and NRA form the basis of developing an early warning system to identify the emergence of any of the top-tier risks, aiming to respond to them before they manifest in a crisis.

Prepare

The emphasis in the preparation stage should be to develop generic or specific capabilities to enable an organization to respond to any situation. The so-called "golden thread" running through this stage is the linkage between the risk registers, the NRA (or the private sector, organizational version of it), and the strategic blend of capabilities required for an effective response. Here, a set of national resilience planning assumptions (NRPA) is required. The NRPA—or a similar set of planning assumptions at the organizational, owner, and operator levels of a critical infrastructure—identifies and predetermines a set of parameters against each risk. These parameters may include, for example, estimates of likely fatalities, casualties, damage estimates, and predictions of how long an organization can tolerate a disruption before it causes serious harm or damage. There are opportunities to include financial data (such as estimated costs due to the loss of life, projected losses caused by disruptions and damage, reputational cost, and risk appetite) against each parameter.

This set of planning assumptions, supported by data in the NRPA (or its equivalent for an organization), informs the characteristics outlined in table 15-2 and will assist in identifying early warning requirements. As mentioned above, the NRPA will also inform capability resource requirements. There are four key elements within this stage: (1) a crisis management plan (CMP), (2) information management and situational awareness plans and processes, (3) crisis management team (CMT) with a clear, accepted structure, composition, and levels of authority, and (4) resilience and capacity within the CMT, including personnel who are well trained, competent, and adequately resourced.

Essential to effective crisis management is the development of an agile, flexible, and current CMP that is concise and well understood by all those who may need to implement it. The CMP should contain key information, to include:

- Authority levels and delegated authorities
- Key CMT staff contact details
- Crisis communications
- Details of how to activate the CMP and CMT
- Information on the organizational crisis response, together with actions to be taken

- Logistical details about where the CMT will convene, its aim and objectives, and logistics support in terms of decision logs, and information management and situational awareness templates

One of the key characteristics of a crisis outlined in table 15-2 is the inherent level of uncertainty that forces responders to operate with ambiguous or insufficient information. This lack of information hinders effective decision making. An imperative during any crisis is to provide policymakers and decisionmakers with as comprehensive an information set as possible or to highlight the difference between (1) what information is known, (2) what information is missing or unknown, and (3) what working assumptions are driving decision making until concrete information is available. Therefore, it is critical to develop information requirements early on to assist decisionmakers and ensure when decisions are made the levels of uncertainty are logged.

Situational awareness is a term that describes the status of knowledge about the event, which may be based upon the team's best efforts during a heightened crisis. To resolve a crisis successfully, it is pivotal to generate and maintain situational awareness and achieve levels of shared situational awareness among those managing the crisis and those responding to it. There may be many barriers to achieving high levels of awareness because of rapidly changing circumstances, the scale of the crisis, technological limitations, and organizational conflicts and differences. In some cases, subject-matter experts may be required to explain data or national security issues may prevent routine sharing of information. It may be prudent to establish an information cell, a situational awareness cell, or a fusion cell staffed by different agencies working collaboratively to coordinate the collection, analysis, and dissemination of information. See chapter 11 for a discussion on information and intelligence sharing best practices.

Response and Recovery

Due to the nature of the crises outlined in table 15-2, it is difficult to prescribe the response to a particular crisis before or while it occurs. Therefore, it is imperative the CMP and CMT are both activated as soon as possible. While there may be some generic first steps or actions that can be implemented right away, the CMT must establish a command center as quickly as possible, generate shared situational awareness, and develop a working strategy. By identifying the issues, making decisions, and assigning or tasking actions, the strategy will be translated into effective

tactical and operational plans. The CMT should record all the decisions made, the actions that were tasked and completed, and key notes from all relevant meetings. As part of the CMP and CMT, an early battle rhythm or operating tempo should be established so that all agencies may report progress and develop shared situational awareness. Paradoxically, the need for quick response and early decision making is often set against a height of uncertainty and lack of information. Delaying decision making and deployment of resources early on in a crisis while waiting for more clarity or certainty to develop, however, will be detrimental to early resolution and recovery.

Strategic planning for an early recovery should be initiated as early as possible. As is often the case during an operation involving multiple government agencies and private-sector representatives, roles and responsibilities change over the duration of managing any incident, emergency, or crisis. Therefore, it is important to include such handover protocols and plans in the CMP.

Crisis Management Team and Leadership

Effective leadership within the CMT is essential during crisis management. As discussed previously, the CMT will have to mobilize quickly and often at short notice. Due to the nature and characteristics of a crisis, the situation does not present ideal conditions in which leaders can make decisions comfortably. Among the many constraints of a crisis, leaders must negotiate a rapidly developing situation, often with an environment characterized by imprecise or unclear information, ambiguity, sometimes chaos, and a fundamental lack of time to balance caution against delay.

In such an environment, leadership within the CMT is crucial to an effective resolution and recovery following a crisis. Leaders and the CMT will face extraordinary pressure and complexity as they tackle the crisis, but a successful resolution depends on timely decision making and good communications. Indeed, these words ring true: Striking the correct balance is critical. Generally, a workable solution delivered on time is more effective than a perfect solution delivered even a little too late.[16]

Training, Exercising, and Learning from Crises

The most effective way to select and develop leaders and the CMT members is through a process of immersive, rigorous, and realistic training and exercising. Organizations should demonstrate a genuine commitment

16. BSI, *Crisis Management*, 16.

to providing the personnel on the CMT with the requisite skills and knowledge to perform their expected duties. The CMP should include a coherent strategy that ensures leaders and the CMT have adequate opportunity to learn, develop, and rehearse the various aspects of the plan. The organization should provide sufficient occasions for repeated training and exercising so the CMT can practice several scenarios, options, and variables to hone their crisis management skills, knowledge, and attitudes. It should be clear that training is for the personnel who serve as leaders and CMT members, while exercising is about rehearsing the plan and interaction with other agencies.

Owners and operators of critical infrastructure, together with those charged with protecting national assets, should appreciate the value of rigorous training and structured exercising in supporting and developing effective crisis management arrangements and plans. Exercises provide a "process to train for, assess, practice, and improve performance in an organization."[17] Effective exercises involve appropriate facilities and entities at local, regional, and national levels to ensure all partners understand, develop, and implement the necessary response plans, including the relevant CMPs. Exercises provide unique opportunities for various response agencies, stakeholders, and local government organizations to validate policies, plans, and procedures while assessing training, equipment, and interorganizational arrangements. This validation and assessment should include systems and processes for developing and implementing information and intelligence-sharing agreements and mechanisms.[18] Intelligence and information sharing among and between partners will assist decision making and ensure effective crisis management by contributing to shared situational awareness.

Properly planned and conducted exercises are a low-cost but high-value investment in relative terms. The process of learning from experience yields many benefits to organizations, and it is essential to incorporate this learning into updated versions of the CMP and renewed guidance to the CMT to ensure continuous improvement. It is also important the owners of the CMP and CMT regularly review post-incident reports and reviews of crises that occur elsewhere. Learning from others' successes, mistakes, and identified areas for improvement will assist an organization in developing and improving the CMP and CMT.

17. BSI, *Societal Security – Guidelines for Exercises* (London: BSI Standards Limited, 2013), 2.

18. BSI, *Societal Security*, 8–9.

NATO and Crisis Management

NATO identifies crisis management as one of its fundamental security tasks, and the 2010 Strategic Concept outlines the Alliance's role in crisis management. As part of its organizational crisis management and response options, NATO can use traditional military means or civil-military cooperation to operate in a myriad of different crises. NATO regularly trains, exercises, and tests its crisis management capability, which it can employ in response to a range of crises initiated by natural or humanitarian disasters or other disruptive events stemming from the political-military realm.

NATO respects the role of sovereign states and does not deploy forces unless it receives political authority to do so. NATO member states and partner nations determine on a case-by-case basis whether to support NATO crisis management operations. NATO also recognizes a military solution alone will not resolve a crisis; rather, a comprehensive suite of political, civilian, and military response options is required to deliver effective crisis management.

Today's international operating environment is evolving, however, and increasing in complexity, with conflict, ungoverned states, protests, pandemics, climate change, extreme weather events, globalization, natural hazards, and emerging threats among these new dynamics. Against this backdrop, NATO is experiencing greater challenges as the apparent shift in the balance of power globally threatens the international rule of law. Therefore, NATO may find opportunities to strengthen the national capability of its Allies and partners to manage crises on a regional or country basis.

Developments in Crisis Management and Resilience

Owners and operators of critical infrastructure face increasingly complex structures, systems, and components that deliver essential services to critical infrastructure systems. Critical infrastructure sectors are becoming more dependent and interdependent on supply chains and other critical infrastructure sectors. See the detailed examination of these dependencies and interdependencies in chapter 12. As the complexity of structures, systems, and components increases not just in their respective sectors but also within the supply chain, the inherent vulnerabilities increase, as does the potential for a single threat to disrupt several critical infrastructure sectors simultaneously. This tight coupling—a key element of normal accident theory that highlights the vulnerabilities and associated frailties arising from interconnected and interdependent technological advances—

is on the rise in most critical infrastructure sectors.[19] Consequently, previous studies involving high-reliability organizations (HRO) are now being revisited by academics and operators of such organizations. HRO theory explains that the conventional understanding of critical infrastructure fails to recognize the complexity of HROs and advocates for a different approach to security and resilience.[20] The following five principles are typically prominent characteristics of HROs:

- Preoccupation with failure. Predominantly highly regulated industries where technical, human, or process failures are addressed, and risk aversion may be high.
- Reluctance to simplify. Some industries are complex by nature, such as aviation, transportation, energy, water, and telecommunications.
- Sensitivity to operations. With a clear focus on what the organization delivers, this front-line view informs decision making.
- Commitment to resilience. The capability and capacity to anticipate the onset of incidents, emergencies, and crises, and to adapt to and overcome them.
- Deference to expertise. Recognizing the importance of expertise rather than authority, they use specialists and experts to tackle uncertainty and rapidly changing circumstances and develop situational assessment on the ground.

Within the context of enhancing CISR posture, capabilities, and policies, these five principles may prove quite helpful to NATO as an organization and to its member states and partner nations. This chapter also highlighted the relevance and utility of the United Kingdom's national standard for crisis management, which Allies and partners may find to be an excellent framework for developing or updating crisis management capability at the local, regional, or national level. Additionally, the chapter advocated the use of exercises to train personnel and validate plans and procedures as a means to enhance CISR in a government or an organization. Allies and partners may find this international standard—published by the International Standards Organization

19. Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Princeton, NJ: Princeton University Press, 1999), 4–6.

20. Karl E. Weick and Kathleen M. Sutcliffe, *Managing the Unexpected: Resilient Performance in an Age of Uncertainty* (San Francisco: Jossey-Bass, 2007).

and used by the United Kingdom and other countries—useful for guidance on planning, conducting, and learning from exercises.[21]

There has also been a shift within nongovernmental sectors, including critical infrastructure and other regulated sectors, to contend with key issues such as risk management principles, crisis management, asset management, information management, assurance and security, business continuity, reputation management, environmental management, and, increasingly, other forms of regulation. Consequently, there has been an emerging trend and interest in combining some of these disciplines together. The drive by the corporate sector and governments to harmonize or synchronize such disparate but connected and interdependent activities has led to an increased interest in the concept of organizational resilience. *Organizational resilience* is defined as the ability to "anticipate, prepare for, and respond and adapt to incremental change and sudden disruptions in order to survive and prosper."[22] At its core, organizational resilience allows an organization to adapt to evolving conditions while retaining its essential values, purpose, and vision. Sometimes, this adaptation means an organization implements response options that it developed prior to a disruption. At other times, it requires the organization to adapt its structures or actions to adjust to new conditions.

Given NATO's June 2021 announcement of the Strengthened Resilience Commitment, there may be an opportunity for the Alliance to integrate its crisis management doctrine with other essential functions and reinforce this commitment by adopting organizational resilience as a new core task. Additionally, NATO also has a central role in civil-military cooperation (CIMIC), and its CIMIC Field Handbook outlines the defense contribution to resilience and assistance to civil authorities. Together, the focus areas of organizational resilience and CIMIC offer potential next steps for NATO to enhance its collective CISR posture.

Summary and Conclusion

In summary, this chapter examined the unique aspects of managing a crisis in contrast to conventional incident and emergency management, and why those traditional approaches will not resolve a crisis effectively. While the first portion of the chapter presented the various stages and phases of crises and the appropriate responses, the following sections discussed the

21. BSI, *Crisis Management*; and BSI, *Societal Security*.

22. "Organizational Resilience," BSI (website), accessed on January 6, 2022, https://www.bsigroup.com/en-US/our-services/Organizational-Resilience/.

key role NATO plays in effective crisis management and the opportunities posed by the Alliance's recent commitment to strengthen resilience. The chapter also identified new initiatives in developing organizational resilience and revisited previous organizational theories relevant to CISR policies and practices. These concepts, particularly organizational resilience, could help NATO in two significant ways. First, they can improve the support NATO as an institution provides to its member states and partner nations when crises arise or emerge. Second, they can assist critical infrastructure stakeholders among the Allies and partners to enhance security and resilience, which benefits national security and resilience and enables NATO to fulfill its core tasks. In conclusion, crises do not remain static nor does the technology that delivers essential services from critical infrastructure. Therefore, the structures, processes, and thinking that underpin CISR and crisis management need to evolve to counter the threats of disruption countries face today.

Given the announcement of the June 2021 resilience commitment and the work toward an updated Strategic Concept commensurate with the current international security environment, there are key CISR and crisis management opportunities for NATO to pursue. Specifically, NATO should review the development of organizational resilience guidance and the existing crisis management standards to support its Strengthened Resilience Commitment. This review would enable benchmarking of current NATO doctrine to identify areas for improvement in the Alliance's core task to support member states and partner countries in responding to and managing crises.

About the Contributors

Chris Anderson is an incident management and infrastructure protection expert with three decades of government, military, and private-sector experience. He is currently the principal adviser for national security and emergency preparedness at Lumen, a US-based global network provider and tech company. He previously held various senior leadership positions in emergency management and national security at the US Federal Communications Commission and US Department of Homeland Security. Anderson began his career as a US Navy helicopter pilot, completing 24 years of active and reserve service. He holds master's degrees in national security strategy from the National War College and in management information systems from Bowie State University, and he received his undergraduate degree from the University of Virginia.

Malcolm Baker has an extensive background in national security, emergency management, incident and crisis management, and critical national infrastructure protection. Baker is currently the director of Resilience Limited in the United Kingdom. He continues to support COE-DAT's Critical Infrastructure Security and Resilience program. He has worked extensively with agencies across the British government, including emergency services, the military, and national security. Previously, Baker was a senior officer with the Counter Terrorism Command and the Metropolitan Police Service Anti-Terrorist Branch. He was awarded a master of science degree from the Defence College of Management and Technology at the Defence Academy of the United Kingdom at Shrivenham. Baker is currently studying to become a certified security management professional and a chartered security professional.

Ronald Bearse is an expert in critical infrastructure protection and national security preparedness, with more than 23 years of experience in the US Departments of Defense, Homeland Security, and Treasury. He is an adjunct professor at the Massachusetts Maritime Academy and an adviser to NATO's Centre of Excellence for the Defence Against Terrorism (COE-DAT), where he teaches in COE-DAT's Critical Infrastructure Protection Against Terrorist Attacks training program. Bearse earned an undergraduate degree in political science and Soviet studies from the University of Massachusetts at Amherst and a master of public administration degree from George Washington University. He is a distinguished graduate of the US National Defense University and a former senior fellow

at George Mason University's Center for Infrastructure Protection and Homeland Security.

Salih Biçakci is an associate professor of international relations at Kadir Has University in Istanbul, Turkey, and a researcher at the university's Center for Cybersecurity and Critical Infrastructure Protection. His research focuses on cybersecurity, critical infrastructure protection, hybrid security, data privacy, and terrorism. Biçakci is an adviser to the NATO Centre of Excellence for the Defence Against Terrorism (COE-DAT) and frequently lectures at COE-DAT and the NATO Centres of Excellence for Command and Control and for Maritime Security. He also teaches courses on cybersecurity and Middle Eastern security at the Turkish War College's Armed Forces Academy. Biçakci has prepared cybersecurity reports for nuclear power plants and led simulation exercises and training for several organizations' executive management.

Steve Bieber has more than 30 years of experience in leading development and reform in water security, public policy, and environmental regulation. He is currently the water resources program director for the Metropolitan Washington Council of Governments (MWCOG) and is responsible for managing its water resources programs, including the regional Anacostia Restoration Partnership, water security programs, drinking water and wastewater planning, drought management, urban stream restoration, and other related environmental programs for local governments and water utilities in the Washington, DC, area. Bieber holds a bachelor of science degree in zoology from Michigan State University, a master of science degree in oceanography from Old Dominion University, and a master of public administration degree from the University of Baltimore.

Sungbaek Cho is currently a researcher in the Strategy Branch at the NATO Cooperative Cyber Defence Centre of Excellence (CCD-COE) in Tallinn, Estonia. The NATO CCD-COE is a multinational and interdisciplinary cyber-defense hub to support NATO and its member nations with unique, interdisciplinary expertise in cyber-defense research, training, and exercises, covering the focus areas of technology, strategy, operations, and law. Cho's research interests include cybersecurity risk management, critical infrastructure protection, national cybersecurity strategy, information sharing, and security certification and accreditation.

Adrian Dwyer is an expert with nearly 40 years of experience in the counterterrorism arena. Adrian served in the British Army Royal Engineers as a bomb disposal officer and an instructor in the Counterterrorist Wing

at the Royal School of Military Engineering. He was appointed subsequently as the principal counterterrorism risk adviser to the British Transport police—a post he held for more than 20 years. Dwyer has also worked as a risk management consultant in a range of industries, from risk reinsurance to petrochemicals. He is a member of the Institution of Royal Engineers and the Institute of Explosives Engineers. He holds a master of science degree in risk, crisis, and disaster management and a PhD from Glasgow University.

Carol V. Evans is director of the Strategic Studies Institute and US Army War College Press at the US Army War College in Carlisle, Pennsylvania. The Strategic Studies Institute is the US Army's leading think tank for geostrategic and national security research and analysis. She brings 30 years of expertise in the areas of mission assurance, crisis and consequence management, asymmetric warfare, terrorism, maritime security, and homeland security. Since 2014, Evans has been a lecturer at NATO's Centre of Excellence for the Defence Against Terrorism (COE-DAT) in Ankara, Turkey, where she teaches in COE-DAT's Critical Infrastructure Protection Against Terrorist Attacks training program. She holds a master of science degree and a doctor of philosophy degree from the London School of Economics.

Geoffrey French has worked in the critical infrastructure protection community since the 1990s. He currently leads the Risk Analysis and Risk Management Cell of the US Cybersecurity and Infrastructure Security Agency's COVID Task Force. Previously, French was the analytic director for security risk at CENTRA Technology, Inc., where he supported numerous programs and designed several risk methodologies for the US Department of Homeland Security, including tools for assessing terrorist risk to infrastructure, security risk to special events, and all-hazards risks to a region. French has also been a lecturer for NATO COE-DAT since 2015. French has a bachelor of arts degree in history from Wichita State University and a master of arts degree in national security studies from Georgetown University.

David Harell has over 40 years of antiterrorism, protective security, and risk management experience in the public and private sectors. He provides consulting and training to governments and major corporations on a wide range of relevant aspects of security topics, including preparedness, critical infrastructure protection, risk management, and crisis response. During his government security career, he served as El Al's regional security manager for Scandinavia and was the commander for the Israeli Security Agency's (ISA) course for aviation security managers. Upon retiring from the ISA after 24 years of service, Harell cofounded and served as managing director

of a leading and renowned international security consulting company. Currently, he lectures at the Berlin School of Economics and Law Master's Program for Security Management.

Alessandro Lazari has been working as a specialist in critical infrastructure protection, resilience, and cyber security since 2004. He is currently a senior key account manager at 24 AG, focused on incident and crisis management in Europe. From 2010–19, he provided policy support to two key initiatives at the European Commission: the European Programme for Critical Infrastructure Protection and Strengthening Europe's Cyber Resilience. Lazari is a fellow in legal informatics at the University of Lecce's School of Law (Italy) and a lecturer at COE-DAT's Protecting Critical Infrastructure Against Terrorist Attacks course. He is the author of *European Critical Infrastructure Protection*, published in 2014 by Springer Inc. He holds a master's degree in law and a PhD in computer engineering, multimedia, and telecommunications.

Raymond Mey has over 35 years of security experience, including 23 years in federal law enforcement and counterterrorism. He served in various leadership roles at the US Federal Bureau of Investigation (FBI), including security for major political and sporting events, crisis management, and hostage rescue. Mey has conducted security vulnerability assessments for numerous Fortune 500 companies and government organizations and developed strategic security plans and training programs to improve and enhance security and safety. He holds a master of arts degree in psychology from Rhode Island College and a bachelor of arts degree in sociology/psychology from the University of Rhode Island. A graduate of the FBI Academy, Mey has been awarded the FBI Medal of Merit and the FBI Shield of Bravery.

Theresa Sabonis-Helf is the chair of the science, technology, and international affairs concentration in the master's degree program at Georgetown University's School of Foreign Service. Previously, she was a professor of national security strategy at the National War College. She has lived and worked in seven countries of the former USSR, assisted two nations in developing their first national security strategies, and coedited two volumes on Central Asia's political and economic transition. Sabonis-Helf has published and lectured extensively on energy security, climate change policies, post-Soviet energy and environmental issues, regional water politics, regional trade and transit, and the politics of electricity. She frequently advises the US Department of State and US Agency for International Development, and is a member of the Council on Foreign Relations.

Duane Verner is the Resilience Assessment Group leader in Argonne National Laboratory's Decision and Infrastructure Sciences Division. Since 2009, he has provided methodology development and project implementation support to the US Department of Homeland Security. Verner is an active member of the European Centre of Excellence for Countering Hybrid Threats and a delegate to the Organisation for Economic Cooperation and Development High Level Risk Forum. In 2018, he was appointed as a civil expert for NATO's Civil Emergency Planning Committee to advise on all aspects of regional energy resilience. Previously, Verner was a project manager for a private-sector engineering firm in New York City, working in the transportation, homeland security, and defense sectors. He is a certified planner with a master of arts degree in urban planning.

The United States Army War College educates and develops leaders for service at the strategic level while advancing knowledge in the global application of Landpower.

The purpose of the United States Army War College is to produce graduates who are skilled critical thinkers and complex problem solvers in the global application of Landpower. Concurrently, it is our duty to the Army to also act as a "think factory" for commanders and civilian leaders at the strategic level worldwide and routinely engage in discourse and debate on the role of ground forces in achieving national security objectives.

The Strategic Studies Institute publishes national security and strategic research and analysis to influence policy debate and bridge the gap between military and academia.

The SSI Live Podcast Series provides access to SSI analyses and scholars on issues related to national security and military strategy with an emphasis on geostrategic analysis. https://ssi.armywarcollege.edu/ssi-live-archive

The Center for Strategic Leadership provides strategic education, ideas, doctrine, and capabilities to the Army, the Joint Force, and the nation. The Army, Joint Force, and national partners recognize the Center for Strategic Leadership as a strategic laboratory that generates and cultivates strategic thought, tests strategic theories, sustains strategic doctrine, educates strategic leaders, and supports strategic decision making.

The School of Strategic Landpower provides support to the US Army War College purpose, mission, vision, and the academic teaching departments through the initiation, coordination, and management of academic-related policy, plans, programs, and procedures, with emphasis on curriculum development, execution, and evaluation; planning and execution of independent and/or interdepartmental academic programs; student and faculty development; and performance of academic-related functions as may be directed by the Commandant.

The US Army Heritage and Education Center makes available contemporary and historical materials related to strategic leadership, the global application of Landpower, and US Army Heritage to inform research, educate an international audience, and honor soldiers, past and present.

The Army Strategic Education Program executes General Officer professional military education for the entire population of Army General Officers across the total force and provides assessments to keep senior leaders informed and to support programmatic change through evidence-based decision making.

US ARMY WAR COLLEGE PRESS

The US Army War College Press supports the US Army War College by publishing monographs and a quarterly academic journal, *Parameters*, focused on geostrategic issues, national security, and Landpower. Press materials are distributed to key strategic leaders in the Army and Department of Defense, the military educational system, Congress, the media, other think tanks and defense institutes, and major colleges and universities. The US Army War College Press serves as a bridge to the wider strategic community.

All US Army Strategic Studies Institute and US Army War College Press publications and podcasts may be downloaded free of charge from the US Army War College website. Hard copies of certain publications may also be obtained through the US Government Bookstore website at https://bookstore.gpo.gov. US Army Strategic Studies Institute and US Army War College publications may be quoted or reprinted in part or in full with permission and appropriate credit given to the US Army Strategic Studies Institute and the US Army War College Press, US Army War College, Carlisle, PA. Contact the US Army Strategic Studies Institute or the US Army War College Press by visiting the websites at: https://ssi.armywarcollege.edu and https://press.armywarcollege.edu.

The US Army War College Press produces two podcast series. Decisive Point, the podcast companion series to the US Army War College Press, features authors discussing the research presented in their articles and publications. Visit the website at: https://ssi.armywarcollege.edu/decisive.

Conversations on Strategy, a Decisive Point podcast subseries. features distinguished authors and contributors who explore timely issues in national security affairs. Visit the website at: https://ssi.armywarcollege.edu/cos.

ISBN 1-58487-840-1
9 781584 878407 90000 >
https://press.armywarcollege.edu