Identifying and Prioritizing Systemically Important Entities

Advancing Critical Infrastructure Security and Resilience

JOHN BORDEAUX, JONATHAN W. WELBURN, SASHA ROMANOSKY, BENJAMIN BOUDREAUX, AARON STRONG, SHANNON PRIER, CHERYL K. MONTEMAYOR, JHACOVA WILLIAMS, JESSICA WELBURN PAIGE, MICHAEL J. D. VERMEER, ZEV WINKELMAN

HSOAC
HOMELAND SECURITY
OPERATIONAL ANALYSIS CENTER

An FFRDC operated by the RAND Corporation under contract with DHS

This research was published in 2023.

Approved for public release; distribution is unlimited.

About This Report

The Cybersecurity and Infrastructure Security Agency's (CISA's) National Risk Management Center (NRMC) has a dual mission: to analyze and reduce risk to the nation's critical infrastructure. Following the recommendations of the Cyberspace Solarium Commission, and a prior Homeland Security Operational Analysis Center (HSOAC) study to explore information technology products and services that could be considered important to protecting critical infrastructure, the NRMC asked that HSOAC: (1) develop methods advancing CISA's concept of risk prioritization through support for the systemically important critical infrastructure (SICI) program initiative;[1] (2) advance prior HSOAC work documented in a previous report,[2] and accompanying materials, and; (3) aid CISA in responding to congressional and Executive Office of the President taskers expected during calendar years 2021 and 2022, and any other relevant changes to CISA requirements. This study was titled "Identifying Systemically Important Critical Infrastructure," although concepts evolved, and the sponsor requested the focus be on identifying and prioritizing systemically important entities.

To address recommendations of the Cyberspace Solarium Commission, CISA requires a more-robust and -sustainable analytic approach to identifying SICI over the long term, and an initial list of candidate entities that might be considered part of the nation's SICI. The framework and continued research regarding critical information technology products and services discussed in this report will enable CISA to improve the cybersecurity and resilience of our nation's infrastructure in the face of systemic risks.

The findings should be of interest to critical infrastructure stakeholders (to include Sector Risk Management Agencies and private sector firms) and researchers and analysts engaged in further definition and operationalization of systemic risk and associated concepts.

This research was sponsored by CISA's NRMC and conducted within the Strategy, Policy, and Operations Program of the HSOAC federally funded research and development center (FFRDC).

Comments or questions on this draft report should be addressed to the project leaders, John Bordeaux at bordeaux@rand.org or Jonathan Welburn at jwelburn@rand.org.

About the Homeland Security Operational Analysis Center

The Homeland Security Act of 2002 (Section 305 of Public Law 107-296, as codified at 6 U.S.C. § 185), authorizes the Secretary of Homeland Security, acting through the Under Sec-

[1] Since renamed within the NRMC as the systemically important entities program.

[2] Sasha Romanosky, John Bordeaux, Michael J. D. Vermeer, Jonathan W. Welburn, Aaron Strong, and Zev Winkelman, *Identifying Critical IT Products and Services*, Homeland Security Operational Analysis Center operated by the RAND Corporation, RR-A923-2, 2022.

retary for Science and Technology, to establish one or more FFRDCs to provide independent analysis of homeland security issues. The RAND Corporation operates HSOAC as an FFRDC for the U.S. Department of Homeland Security (DHS) under contract HSHQDC-16-D-00007.

The HSOAC FFRDC provides the government with independent and objective analyses and advice in core areas important to the department in support of policy development, decisionmaking, alternative approaches, and new ideas on issues of significance. The HSOAC FFRDC also works with and supports other federal, state, local, tribal, and public- and private-sector organizations that make up the homeland security enterprise. The HSOAC FFRDC's research is undertaken by mutual consent with DHS and is organized as a set of discrete tasks. This report presents the results of research and analysis conducted under 70RCSA21FR0000046, Identifying Systemically Important Critical Infrastructure.

The results presented in this report do not necessarily reflect official DHS opinion or policy.

For more information on HSOAC, see www.rand.org/hsoac.

Acknowledgments

We are grateful to Linda Ward, Susan Stevens, Jay Healey, Henry Willis, and Stefan Haus for their enduring support, feedback, and advice. Furthermore, we are thankful to Lloyd Dixon, Quentin Hodgson, Peter Schirmer, and Craig Bond for their continued reviews of our work.

Summary

Issue

Against the backdrop of increasingly impactful cyber threats, the John S. McCain National Defense Authorization Act for Fiscal Year 2019 established the Cyberspace Solarium Commission (henceforth Solarium) to "develop a consensus on a strategic approach to defending the United States in cyberspace against cyberattacks of significant consequences."[3] Solarium called for a strategy of *layered cyber deterrence* that (1) promotes responsible behavior that works against malicious activity in cyberspace, (2) denies benefits to cyber adversaries, and (3) imposes costs against cyber adversaries who engage in malicious activity. A key recommendation in the report was that Congress "codify the concept of *systemically important critical infrastructure*, whereby entities responsible for systems and assets that underpin national critical functions [NCFs] are ensured the full support of the U.S. government and shoulder additional security requirements befitting their unique status and importance."[4]

Following the recommendations of the Solarium report, the National Risk Management Center (NRMC) (on behalf of the U.S. Department of Homeland Security's [DHS's] Cybersecurity and Infrastructure Security Agency [CISA]) was asked to organize and design a program to identify systemically important critical infrastructure (SICI) entities—subsequently relabeled as systemically important entities (SIEs)—and to develop a plan to better anticipate and address potential similar incidents in the future. The NRMC asked the Homeland Security Operational Analysis Center (HSOAC), a federally funded research and development center operated by the RAND Corporation, to conduct analyses to inform these efforts.

Approach

First, we developed a working framework to identify entities that are systemically important to critical infrastructure. We then introduced methodology for prioritizing potential SIEs based on their size and interconnectedness. For both types of metrics, we provide approaches for data-driven thresholds that could provide shorter lists of prioritized SIEs. We then built on existing work regarding critical information technology (IT) products and services to extend the analysis to federal agencies and firms that install potentially vulnerable software in addition to those that write the software. We used extensive data management, data analytics, and visualization to understand dependencies across software and associated librar-

[3] Public Law 115-232, John S. McCain National Defense Authorization Act for Fiscal Year 2019, 2018.

[4] U.S. Cyberspace Solarium Commission, *Cyberspace Solarium Commission Report*, Washington, D.C., March 2020.

ies. We then employed data integration and data management methods to develop a Tableau visualization dashboard to process initial lists of entities.

Key Outcomes

Definitions of Systemic Importance and Systemically Important Entities

We examined previous discussions of systemic risk drawn from historical observations of economic crises and examinations of systemic cyber risk to develop the following definitions of *systemic importance* and *SIEs*:

- *Systemic importance*: A condition derived from the characteristics of an infrastructure, resource, component, or entity, such as its size, interconnectedness, substitutability, and complexity that increases the potential impact of its disruption or failure on others within the system or beyond, to affect to a debilitating degree national security, national economic security, public health, or public safety, or any combination thereof. Examples include a physical infrastructure; an asset in a virtual network; a component in technology; or an organization, the resources they depend on, and services they provide.
- *SIEs*: Entities that own, operate, or otherwise control critical infrastructure and are prioritized by the CISA director as systemically important based on the potential impact that their destruction or incapacity (to include disruption, corruption, or dysfunction) will have a debilitating systemic or cascading impact on NCFs, national security, national economic security, or public health or safety. CISA is dealing with entities because they are able to participate as risk managers in reducing national systemic risk.

Approach for Identifying Systemically Important Entities

We used these definitions in developing a transparent, data-driven methodology for identifying and prioritizing SIEs. The approach first identifies SIEs as entities that have the potential to disrupt one or many NCFs by determining which entities draw revenue from at least one NCF. Specifically, we introduce a two-step process of (1) connecting NCFs to economic sectors and (2) connecting economic sectors to specific entities. This approach provides a long list of thousands of entities with the potential to become SIEs.

Given the sponsor requirement to provide manageable lists of 100 or 250 SIEs, the approach prioritizes SIEs based on their **size** and **interconnectedness**. For size, we estimate which entities appear to be important economically, based on their overall revenue and their revenues drawn from NCF-specific business lines. For interconnectedness, we estimate an entity's centrality in a large interfirm network of customer and supplier relationships (i.e., supply chain networks). Additionally, we identified the limitations of a largely economic approach and highlighted the potential for considering the role of equity in an entity's systemic importance.

Analytic Platform to Process Initial Lists of Entities Associated with National Critical Functions

We also developed an analytic platform, the Systemic Importance Analytic Model (SIAM), to process initial lists of entities associated with NCFs. We used data analytics and visualization to connect and visualize data from multiple commercial data vendors, revisiting and augmenting those used in the earlier study of critical IT products and services. We employed data integration and data management methods and developed SIAM using the Tableau visualization dashboard. This tool allows an analyst to understand and rank SIEs based on several factors, as discussed further in Chapter 3 and Appendix A.

Real-Time Analytic Support

Finally, part of our task was to help the NRMC adapt to an evolving set of mission-related tasks associated with expected executive orders, draft legislation, and a new administration's vision for CISA and the NRMC. We provided real-time analytic support for the NRMC's response for Executive Order 14028 (2021) and ongoing coordination with the National Cyber director and the White House National Security Council's Senior Director for Resilience and Response.

Conclusion

This report does not include the lists of prioritized firms associated with NCFs, because those lists will be revisited annually. Instead, this report focuses on the "how" and offers approaches and future analytic initiatives to further refine the NRMC's efforts to understand and mitigate systemic risk to the nation's critical infrastructure. Therefore, the results of this report provide the NRMC with potential objective criteria for determining a prioritized list of SIEs—a list which can enable CISA to strengthen entity risk management and coordination, allocate resources, monitor threats and hazards, and prioritize planning in support of a broader national strategy of layered deterrence.

Significant work remains in developing concepts and modeling approaches for systemic risk to critical infrastructure, advancing the NRMC's incorporation and stewardship of data sets for analysis and visualization, maturing the SIE Program Office processes and procedures for analysis and outreach, and advancing SIAM to reflect emerging perspectives for prioritization—including public health and safety, national security, equity, and other areas.

We identify several analysis needs that would help advance the NRMC's risk reduction mission. These include: (1) advancing SIE concepts and modeling approaches; (2) developing data management methods and planning for analytic input data; (3) advancing SIE as a sustainable program; and (4) refining the SIE analytic platform.

Contents

About This Report ... iii
Summary ... v
Figures and Tables ... xi

CHAPTER ONE
Introduction ... 1
Focus of This Study ... 3
Development of an Analytic Framework for Identifying Systematically Important Entities ... 4
Organization of This Report ... 5

CHAPTER TWO
Background and Motivation on Systemic Importance ... 7
Lessons from Systemically Important Financial Institutions ... 7
Lessons from Systemic Risk ... 9
Systemically Important Entities ... 12

CHAPTER THREE
Systemically Important Entities: Identification and Prioritization ... 17
Identification of Systemically Important Entities ... 17
Prioritization of Systemically Important Entities ... 25
Limitations ... 28
Overcoming Limitations from an Equity Perspective ... 31

CHAPTER FOUR
Systemic Cyber Risk ... 39
Cyber Risk of Systemically Important Entities ... 40
Cyber Risks of the Software Supply Chain ... 45
Limitations in Current Approaches to Understanding Systemic Cyber Risk That Might Be Addressed Through Future Work ... 51

CHAPTER FIVE
Future Research ... 55
Advance Systematically Important Entity Concepts and Modeling ... 55
Develop Data Management Methods and Plan for Analytic Input Data ... 56
Advance SIE Program as a Sustainable Program ... 57
Refine the Systemic Importance Analytic Model ... 57

APPENDIXES

A. Systemic Importance Analytic Model 59
B. Cybersecurity and Infrastructure Security Agency Strategic Intent and National Risk Management Center Missions and Objectives 63
C. Cyber Data and Software Dependencies 65
D. Analytical Support for the Cybersecurity and Infrastructure Security Agency 69

Abbreviations 79
References 81

Figures and Tables

Figures

3.1. Network Centrality 26

3.2. Knee-in-the-Curve Graph for Fuel Refining and Processing. 27

3.3. Distribution of Entity Types Across National Critical Functions. 30

4.1. Cyberattack Surface 43

4.2. Identify Software and IT Companies 44

4.3. Software and IT Supply Chain for List of Systemically Important Entities 45

4.4. Illustration of Software Dependents Tree 47

4.5. Illustration of a Software Bill of Materials Dependency Tree 48

4.6. Mapping Software Supply Chain 50

A.1. Systemic Importance Analytic Model Dashboard 60

C.1. Dependency Network of Apache-Log4J 66

D.1. Finding the Knee-in-the-Curve for the Provide Medical Care National Critical Function 77

Tables

2.1. The 55 National Critical Functions 15

3.1. National Critical Function to North American Industry Classification System Mapping 19

3.2. Share of Entities by Entity Type 29

4.1. Advantages and Disadvantages of Custom Versus Commercial Cyber Risk Scores 41

4.2. Systemically Important Entity Cyber Risk Approaches 46

4.3. Software Cyber Risk Approaches. 49

C.1. Cyber Data Sources. 65

C.2. Apache-Log4J Dependencies 67

D.1. Top Five Products by Market Share and Alignment to Federal Wallet Share Category. 72

D.2. Natonal Critical Function Analytic Phases 74

CHAPTER ONE

Introduction

Years before SolarWinds there was Titan Rain. In 2003, Chinese state hackers gained access to sensitive information and information systems held by the U.S. Department of Defense (DoD), the Federal Bureau of Investigation, Lockheed Martin, the National Aeronautics and Space Administration (NASA), Sandia National Labs, and others.[1] Titan Rain was one of the first notable cyber exploits used against the United States, and it was also a warning for attacks to come.

More recently, cyber actors have launched attacks against national critical functions (NCFs) and critical infrastructure. From 2012 to 2013, Iranian hackers breached the systems of several major financial institutions, including JP Morgan Chase, and a flood control dam in upstate New York. In 2014, Chinese hackers breached the unclassified computing networks of the U.S. Department of State and, in 2015, gained access to 21.5 million sensitive records from the Office of Personnel Management. In 2016, Russian state actors (1) breached the systems of the Democratic National Committee with the intention of hindering the NCF of conducting elections, (2) successfully disrupted power to 225,000 Ukrainian households through cyber means, and (3) quietly gained access to assets on the U.S. power grid.[2] In 2017, malware attacks exploiting the Windows EternalBlue vulnerability known as WannaCry and NotPetya quickly spread around the globe, leading to large disruptions across sectors, particularly in health care and education, while driving large economic losses:[3] The U.S. pharmaceutical firm Merck recorded $1.4 billion in total costs from the NotPetya incident alone.[4]

Against this overall backdrop of increasing danger in cyberspace, Section 1652 of the John S. McCain National Defense Authorization Act for Fiscal Year 2019 established the "Cyberspace Solarium Commission" to "develop a consensus on a strategic approach to

[1] Nathan Thornburgh, "The Invasion of the Chinese Cyberspies," *Time*, Washington, D.C., 2005; Richard Norton-Taylor, "Titan Rain: How Chinese Hackers Targeted Whitehall," *The Guardian*, Vol. 4, 2007.

[2] Nicole Perlroth and David E. Sanger, "Cyberattacks Put Russian Fingers on the Switch at Power Plants, U.S. Says," *New York Times*, March 15, 2018.

[3] Kim S. Nash, Sara Castellanos, and Adam Janofsky, "One Year After NotPetya Cyberattack, Firms Wrestle With Recovery Costs," *Wall Street Journal*, June 27, 2018.

[4] Andrea Vittorio, "Merck's $1.4 Billion Insurance Win Splits Cyber From 'Act of War'," *Bloomberg Law*, January 19, 2022.

defending the United States in cyberspace against cyberattacks of significant consequences."[5] When the Cyberspace Solarium Commission (henceforth Solarium), co-chaired by Senator Angus King and Representative Mike Gallagher, released its report in 2020, it called for a strategy of *layered cyber deterrence* that (1) shapes behavior against malicious activity in cyberspace, (2) denies benefits to cyber adversaries who seek it, and (3) imposes costs against cyber adversaries who engage in it. The Solarium report made more than 80 recommendations grouped under six pillars: "(1) reform the U.S. government's structure and organization for cyberspace; (2) strengthen norms and nonmilitary tools; (3) promote national resilience; (4) reshape the cyber ecosystem; (5) operationalize cybersecurity collaboration with the private sector; and (6) preserve and employ the military instrument of national power."[6] However, to accomplish these objectives, particularly pillar five, the Solarium report acknowledged a key gap in defending systemically important critical infrastructure (SICI):

> The U.S. government should improve government support to private-sector cyber defensive operations. However, the federal government has limited resources and capabilities, and should prioritize the defense of systemically important critical infrastructure—the critical infrastructure entities that manage systems and assets whose disruption could have cascading, destabilizing effects on U.S. national security, economic security, or public health and safety. While the U.S. government has taken steps to assist these high-risk entities through Section 9 of Executive Order 13636, that effort falls short of codifying or fully implementing the social contract of shared responsibility and partnership in cybersecurity—and it also does not empower the U.S. government with the resources and authorities necessary to defend them.

To address this gap, the Solarium report recommended (5.1) that the concept of "systemically important critical infrastructure" needed to be clearly defined and codified:

> Congress should codify the concept of "systemically important critical infrastructure", whereby entities responsible for systems and assets that underpin national critical functions are ensured the full support of the U.S. government and shoulder additional security requirements befitting their unique status and importance.

For Solarium, identifying SICI was essential to the strategy of layered cyber deterrence to understand both how to deny benefits and when to impose costs in defense of SICI. More specifically, the Solarium report stated that Congress had a responsibility to provide resources

[5] Public Law 115-232, 2018.

[6] U.S. Cyberspace Solarium Commission, *Cyberspace Solarium Commission Report*, Washington, D.C., March 2020.

to the U.S. Department of Homeland Security's (DHS's) Cybersecurity and Infrastructure Security Agency (CISA) to support the agency's efforts in identifying SICI:

> Congress should recognize and provide sufficient resources to support CISA's emergent efforts to identify and mitigate risks to national critical functions and to serve as the primary federal entity responsible for organizing and coordinating whole-of-government, public-private activities to identify, assess, and manage national risk. As detailed in the "Promote National Resilience" pillar, Congress should codify CISA responsibilities and ensure sufficient resources for its national risk management programs, including its support to sector-specific agencies, its critical role in Continuity of the Economy planning (recommendation 3.2), and its identification of systemically important critical infrastructure (recommendation 5.1).

These recommendations provide the core motivation for this report. Following the recommendations of the Solarium report, CISA's National Risk Management Center (NRMC) was tasked with developing an initiative to identify SICI.

Focus of This Study

The NRMC was asked to organize and design a program to identify SICI entities (Solarium recommendation 5.1), and develop a plan to better anticipate and address potential similar incidents in the future. The NRMC asked the Homeland Security Operational Analysis Center (HSOAC), a federally funded research and development center operated by the RAND Corporation, to conduct analysis to inform these efforts.

A key objective of the analysis was to identify SICI. The 2020 revelation of the widespread SolarWinds supply chain computer network exploitation highlighted both the systemic risk posed by new threats to key firms and infrastructure and the challenge in identifying SICI. SolarWinds brought several questions to the fore, among them the following: What are today's sources of systemic risk, and can we develop an understanding of software (and possibly hardware) vulnerabilities and criticalities to prevent an exploitation or shorten the recovery cycle for those that occur?

During the study, the scope was redefined as identifying and prioritizing *systemically important entities* (SIEs), initially defined as private sector firms deemed central to certain NCFs through economic and network analytic methods to determine dependencies between firms. Although the specter of systemic cyber risk provided the primary motivation for identifying SIEs, the concept encompasses several risks beyond cyber risks. These vary from the traditional focus on systemic financial risks to emerging sources of systemic risk (e.g., climate shifts, conflict events, and pandemic disruptions).

The identification and prioritization of SIEs can help the NRMC prepare for, and respond to, a variety of future threats and hazards. Furthermore, the 2022 crisis in Europe accelerated interest in understanding systemic risk to U.S. critical infrastructure within the national security community and Congress. The resulting dynamics required an agile response to dynamic tasking from the NRMC sponsor throughout the course of this study.

Development of an Analytic Framework for Identifying Systemically Important Entities

We developed a working framework to identify entities that are systemically important to critical infrastructure. This analytic framework was intended to be capable of defensibly identifying SIEs and elucidating authorities for engaging these entities and stakeholders for reducing systemic risk. The framework was developed in close coordination with the DHS Office of the Chief Economist (OCE), the NRMC Analytic Division, and other CISA stakeholders.

The resulting framework is traceable and transparent, and represents the first data-driven representation of SIEs. Prior to this study, the only way to develop these lists was through a consensus model that included direct coordination with Sector Risk Management Agencies as the first step. The development of a transparent analytic framework that produces initial data-driven lists of SIEs provides the NRMC with a consistent approach that will aid Sector Risk Management Agencies in further refining the results; there is no longer the need to develop them from a blank page.

In addition, we built on existing work regarding Critical Information Technology (IT) Products and Services (CITPS) to extend the analysis to both federal agencies and firms that **install** potentially vulnerable software and to those that **write** the software. Given the additional data sources now available from commercial data providers, this task required extensive data management, data analytics, and visualization to understand dependencies across software and associated libraries.

This analytic support led, unexpectedly, to the development of the Systemic Importance Analytic Model (SIAM), which uses the Tableau visualization dashboard to process initial lists of entities associated with NCFs. The tool allows analysts to understand and rank SIEs based on several factors. We developed the tool using data analytics and visualization to connect and visualize data from multiple commercial data vendors, revisiting those used in the original CITPS study and augmenting them with new sources. We then employed data integration and data management methods to develop the dashboard.

Finally, we were asked to leverage the platform to assist with the NRMC's outreach regarding the SIE lists. This included providing familiarization training and adjudicating questions regarding the draft lists, focusing on adjudicating areas associated with the platform's logic and methods.

Focus of This Report

This capstone report documents the methodologies developed and those currently being explored to identify and prioritize SIEs based on several perspectives beyond economic and network analyses. The lists of prioritized firms associated with NCFs will not be included here, because the lists will be revisited annually. Rather, this report focuses on the "how" and offers approaches and future analytic initiatives to further refine DHS' efforts to understand and mitigate systemic risk to the nation's critical infrastructure.

Organization of This Report

The remainder of this report is organized as follows:

- Chapter Two provides additional background and motivation for the study and provides a definition of SIEs.
- Chapter Three describes the approach we developed to identify and prioritize SIEs. It also addresses the limitations of the approach and options for overcoming those limitations from an equity perspective.
- Chapter Four discusses data-driven approaches for identifying and prioritizing cyber risk and identifies areas for future research.
- Chapter Five describes directions for future research that advance SIEs concepts and modeling, the need to develop data management plan, and future directions for refining the SIE Analytic Platform.

The report also contains four appendixes. Appendix A describes the SIAM. Appendix B describes CISA's strategic intent and the NRMC's missions and objectives. Appendix C discusses cyber data and software dependences. Finally, Appendix D summarizes the tasking and insights developed as part of our work for the NRMC.

CHAPTER TWO

Background and Motivation on Systemic Importance

In this chapter we provide additional background on the motivation for this study, focusing on lessons learned from systemically important financial institutions (SIFI). We also provide a definition of SIEs used in this analysis.

Lessons from Systemically Important Financial Institutions

A decade before Solarium recommended designating SICI, financial reform designated SIFIs. In 2010, Congress passed the Dodd-Frank Wall Street Reform and Consumer Protection Act (henceforth, Dodd-Frank), which aimed to reign in systemic risk posed by financial institutions that were either too big or too interconnected to fail without significant impacts on the financial system. Across 845 pages, 16 titles, and 225 new rules, Dodd-Frank was the most significant act of financial reform in the United States since the passage of the Glass-Steagall Act in 1933.[1] It established the Financial Stability Oversight Council (FSOC), which was given the power to designate nonbank financial companies and financial market utilities as systemically important.[2]

Section 113(a) of Dodd-Frank outlined the characteristics of observable failure points where the FSOC could designate a nonbank financial company as systemically important if they determined that "material financial distress at the U.S. nonbank financial company, or the nature, scope, size, scale, concentration, interconnectedness, or mix of the activities of the U.S. nonbank financial company, could pose a threat to the financial stability of the United States."

[1] V. V. Acharya and M. Richardson, "Implications of the Dodd-Frank Act," *Annual Review of Financial Economics*, Vol. 4, No. 1, 2012.

[2] Although the term *systemically important financial institution* does not appear in Dodd-Frank, it has become synonymous with the "large, inter-connected bank holding companies or nonbank financial companies" Dodd-Frank established the FSOC to monitor for risks to financial stability (Public Law 111-203, Dodd-Frank Wall Street Reform and Consumer Protection Act, 2010, § 112[a]).

Definitions of Systemically Important and Systemic Importance

To designate *systemically important* financial market utilities, Section 804 provided that FSOC "shall designate those financial market utilities or payment, clearing, or settlement activities that the Council determines are, or are likely to become, systemically important." Section 803 provided the following supporting definitions:

> SYSTEMICALLY IMPORTANT AND SYSTEMIC IMPORTANCE—The terms "systemically important" and "systemic importance" mean a situation where the failure of or a disruption to the functioning of a financial market utility or the conduct of a payment, clearing, or settlement activity could create, or increase, the risk of significant liquidity or credit problems spreading among financial institutions or markets and thereby threaten the stability of the financial system of the United States.

International Policy Reforms Concerning Systemically Important Financial Institutions

International policy reforms concerning SIFIs were developed in parallel with the criteria set forth by Dodd-Frank and operationalized by FSOC.[3] In 2009, the International Monetary Fund, Bank for International Settlements, and Financial Stability Board (FSB) reported initial findings on identifying SIFIs,[4] while the Basel Committee on Banking Supervision (BCBS) began deliberating reforms known as Basel III to strengthen global systemically important banks (G-SIBs) against failure.[5] By November 2011, and in coordination with FSB, BCBS published a G-SIB assessment methodology that provided measurable characteristics of systemic importance: size, interconnectedness, substitutability, complexity, and cross-jurisdictional activity.[6]

Motivations for Creating Systemically Important Financial Institutions

Similar to the call for SICI, SIFIs were defined as a response to growing systemic risk caused by the complexity, size, and interconnectedness of financial institutions. However, although cyber threats to SICI have occurred as isolated incidents, the creation of SIFIs was motivated by the large systemic failures that materialized during the 2008 global financial crisis. The

[3] Daniel E. Nolle, "U.S. Domestic and International Financial Reform Policy: Are G20 Commitments and the Dodd-Frank Act in Sync?" Board of Governors of the Federal Reserve System, International Finance Discussion Papers, No. 1024, July 2011.

[4] Financial Stability Board, International Monetary Fund, and Bank for International Settlements, *Guidance to Assess the Systemic Importance of Financial Institutions, Markets, and Instruments: Initial Considerations—Background Paper*, October 28, 2009.

[5] Bank for International Settlements, "Global Systemically Important Banks: Assessment Methodology and the Additional Loss Absorbency Requirement," webpage, November 23, 2021.

[6] Bank for International Settlements, 2021; Bank for International Settlements, "The G-SIB Assessment Methodology—Score Calculation," Basel Committee on Banking Supervision, 2014.

interconnectedness of risks in the financial sector coupled with a poor understanding of the risk landscape led to systemic risk that was large enough that a few entities could lead to cascading failures in the financial sector. By late 2008, the Dow Jones Industrial Average had suffered all-time losses, the economy had lost almost a quarter of a million jobs in a single month, major investment banks ceased to exist, and economic activity declined in more than half of the world's countries.[7] In the aftermath of the financial crisis, Alan Greenspan asked what many regulators and financial experts were wondering: "How did so many experts, including me, fail to see it coming?"[8]

Conclusions

A primary lesson from 2008 was that the interconnections between risks prove to be more consequential than a singular hazard. Dodd-Frank and its global counterpart policy reforms aimed to learn from the financial crisis and enhance response-to-risk signals.

However, although SIFI is narrowly focused on institutions important to the functioning of the financial system, the designation of SICI is motivated by a rising tide of cyber and a worrying new source of systemic risk to NCFs.

Similar to the call for SICI, SIFIs were created as a response to growing systemic risk caused by entities whose failure could disrupt a broader system. However, although the designation of SIFI was motivated by large systemic failures and the resulting systemic impacts on the financial system, the designation of SICI has been motivated by a rising tide of cyber and a worrying new source of systemic risk to NCFs.

Lessons from Systemic Risk

For the most part, the concept of *systemic risk* is drawn from historical observations of economic crises. In a 1995 speech on risk measurement and systemic risk, Alan Greenspan described the difficulty inherent to explaining systemic risk:

> [i]t is generally agreed that systemic risk represents a propensity for some sort of financial system disruption. Nevertheless, after the fact, one observer might use the term 'market failure' to describe what another would deem to have been a market outcome that was natural and healthy, even if harsh. Even with agreement on what constituted a realization of a systemic crisis in financial markets, descriptions of the symptoms of systemic risk cannot be disentangled from theories of how financial crises come to pass. Until we have

[7] Alan Greenspan, "Never Saw it Coming: Why the Financial Crisis Took Economists by Surprise," *Foreign Affairs*, Vol. 92, No. 6, November/December 2013, pp. 88–96.

[8] Greenspan, 2013.

> a common theoretical paradigm for the causes of systemic stress, any consensus on how to measure systemic risk will be difficult to achieve.[9]

This point is more succinctly described by Bisias et al., in the 2012 publication from the U.S. Treasury Department's Office of Financial Research established by Dodd-Frank:

> Systemic risk may be hard to define but they [policymakers] know it when they see it, such a vague and subjective approach is not particularly useful for measurement and analysis, a prerequisite for addressing threats to financial stability.[10]

Definitions of Systemic Risk

Over the years, numerous efforts have been made to define *systemic risk*. For example, De Bandt and Hartmann (2000) define it as the potential to result in "an event, where the release of 'bad news' about a financial institution, or even its failure, or the crash of a financial market leads in a sequential fashion to considerable adverse effects on one or several other financial institutions or markets, e.g., their failure or crash."[11] More recently, the European Central Bank defined systemic risk as "a risk of financial instability so widespread that it impairs the functioning of a financial system to the point where economic growth and welfare suffer materially."[12] Schwarcz notes that a common factor is a "trigger event, such as an economic shock or institutional failure, [that] causes a chain of bad economic consequences-sometimes referred to as a domino effect."[13] Schwarcz observed that systemic risk does not necessarily have to result in outright failure, but could lead to significant losses and substantial volatility and also notes that systemic risk poses a unique challenge by undercutting the risk management strategies of modern portfolio theory, relying on diversification across negatively (or un-) correlated assets through the positive correlations of systemic event and market disruption.

[9] Alan Greenspan, *Remarks at a Research Conference on Risk Measurement and Systemic Risk, Statements and Speeches of Alan Greenspan*, Washington, D.C.: U.S. Federal Reserve, 1995.

[10] Dimitrios Bisias, Mark Flood, Andrew W. Lo, and Stavros Valavanis, "A Survey of Systemic Risk Analytics Working Paper," Office of Financial Research, working paper, No. 0001, 2012.

[11] Olivier De Bandt and Philipp Hartmann, "Systemic Risk: A Survey," European Central Bank, working paper, No. 35, November 2000.

[12] European Central Bank, Financial Stability Review, European Central Bank, IV Special Features, 2010.

[13] Steven L. Schwarcz, "Systemic Risk," *Georgetown Law Journal*, Vol. 97, No. 1, 2008.

Systemic Risk in the Context of Cyber Risk

However, the sources of systemic risk extend far beyond the world of finance, particularly as cyber risk is increasingly viewed as a source of systemic risk. The World Economic Forum (2016) defines *systemic risk* in the context of cyber risks as follows:

> Systemic cyber risk is the risk that a cyber event (attack(s) or other adverse event(s)) at an individual component of a critical infrastructure ecosystem will cause significant delay, denial, breakdown, disruption or loss, such that services are impacted not only in the originating component but consequences also cascade into related (logically and/or geographically) ecosystem components, resulting in significant adverse effects to public health or safety, economic security or national security. The adverse real economic, safety and security effects from realized systemic risk are generally seen as arising from significant disruptions to the trust in or certainty about services and/or critical data (i.e., the integrity of data), the disruption of operations and, potentially, the incapacitation or destruction of physical assets.[14]

The Homeland Security Systems Engineering and Development Institute has defined *systemic cyber risk* according to 11 factors: common mode/repeated attacks, common mode/scattershot attacks, common mode/pervasive attacks, rolling attacks, transitive attacks, cascading attacks, shared resource consumption attacks, critical function attacks, regional attacks, service dependency attacks, and coordinated supply chain attacks.[15]

More succinctly, Welburn and Strong (2021) define three categories of systemic cyber risk—cascading, common cause, and independent:

- "[C]ascading cyber failures are the result of one cyber incident propagating outward and causing many disruptions"
- "[C]ommon cause cyber failures are the result of one cyber exploit triggered at many firms causing many cyber incidents"
- "[I]ndependent cyber failures are the result of cyber incidents exploiting independent vulnerabilities at individual firms and organizations."[16]

Systemic risks, of course, could exist in many domains beyond finance or cyber. The Systemic Risk Centre of the London School of Economics defines *systemic risk* more broadly as "the risk of a breakdown of an entire system rather than simply the failure of individual parts. In a financial context, it captures the risk of a cascading failure in the financial sector, caused

[14] World Economic Forum, "Understanding Systemic Cyber Risk," Global Agenda Council on Risk and Resilience, White Paper, October 2016.

[15] Deborah J. Bodeau and Catherine D. McCollum, *System-of-Systems Threat Model*, Homeland Security Systems Engineering and Development Institute, 2018.

[16] Jonathan W. Welburn and Aaron Strong, "Systemic Cyber Risk and Aggregate Impacts," *Risk Analysis*, 2021.

by interlinkages within the financial system, resulting in a severe economic downturn." Even more broadly, Schweizer (2019) characterizes systemic risk according to the five factors of "complexity and interdependency, transboundariness, nonlinearity, tipping points, and lag in regulation and perception."[17]

However, the World Economic Forum (2021) might provide the simplest and clearest definition of *systemic risk*:

> Seemingly isolated risks that grow and spread across heavily interconnected and deeply ingrained products, services and systems over a defined time horizon. Upon inception, this type of risk cannot be resolved by a single entity or through the broader diversification of organizational operations.[18]

Systemically Important Entities

During the course of this study, the terminology of interest shifted from SICI to systemically important entities (SIE). This reflects a refinement in scope to focus on the firms that own the assets associated with the provision and management of NCFs. However, the underlying intent of this study, to identify critical infrastructure entities that would cause severe harm to the United States if compromised, has remained the same.

Systemically Important Critical Infrastructure

The concept of SICI builds on previous definitions of critical infrastructure. The Critical Infrastructures Protections Act of 2001 defines *critical infrastructure* as:

> [S]ystems 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.[19]

Of these systems and assets, Solarium recommends identifying the entities responsible for these assets so that "entities responsible for systems and assets that underpin national critical functions are ensured the full support of the U.S. government and shoulder additional secu-

[17] Pia-Johanna Schweizer, "Systemic Risks—Concepts and Challenges for Risk Governance," *Journal of Risk Research*, Vol. 24, No. 1, 2019.

[18] World Economic Forum, "Beneath the Surface: Technology-Driven Systemic Risks and the Continued Need for Innovation," Future of Financial Services Series, October 28, 2021.

[19] U.S. Code Title 42, Section 5195c, Critical Infrastructures Protections Act of 2001.

rity requirements befitting their unique status and importance."[20] The Securing Systemically Important Critical Infrastructure Act, as originally introduced, states that an element of critical infrastructure shall be designated as systemically important under these conditions:

- The likelihood that a disruption to, or compromise of, such element of critical infrastructure would result in a debilitating effect on national security, economic security, public health or safety, or any combination thereof.
- The extent to which damage, disruption, or unauthorized access to such element or collectively to the category of critical infrastructure to which such element belongs (i) would disrupt the reliable operation of a category of critical infrastructure; and (ii) would impede provisioning or a national critical function.
- The extent to which increasing the risk management coordination between the Federal Government and the owner or operator of the element would enhance the cybersecurity resiliency of the United States.[21]

The NCFs referred to in these definitions are the 55 "functions of government and the private sector so vital to the United States that their disruption, corruption, or dysfunction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof."[22] CISA organizes the 55 NCFs into four categories: Connect, Distribute, Manage, and Supply. The full set of NCFs are reproduced in Table 2.1.

A Definition of Systemically Important Entities

Although we adopt the definition of *systemic risk* from World Economic Forum (2021) as "[s]eemingly isolated risks that grow and spread across heavily interconnected and deeply ingrained products, services and systems over a defined time horizon,"[23] we leverage the previously mentioned literature and policy discussions of systemic risk, SIFI, SICI, and critical infrastructure to put forward the following new definitions of *systemic importance* and *systemically important entities*:

- **Systemic importance.** A condition derived from the characteristics of an infrastructure, resource, component, or entity (such as its size, interconnectedness, substitutability, and complexity) that increases the potential impact that its disruption or failure could have on others within the system or beyond to affect to a debilitating degree (1) national security, (2) national economic security, (3) public health, (4) public safety, or (5) any combination thereof. Examples include a physical infrastructure, an asset in a virtual

[20] U.S. Cyberspace Solarium Commission, 2020.

[21] U.S. House of Representatives, "Securing Systemically Important Critical Infrastructure Act," Bill 5491, 117th Congress, October 5, 2021.

[22] CISA, "National Critical Functions Set," fact sheet, April 2019.

[23] World Economic Forum, 2021.

network, component in technology, an organization, the resources they depend on, and services they provide.

- **SIEs.** Entities that own, operate, or otherwise control critical infrastructure and are prioritized by the CISA director as systemically important based on the potential impact that their destruction or incapacity (to include disruption, corruption, or dysfunction) will have a debilitating systemic or cascading impact on NCFs, national security, national economic security, or public health or safety. CISA is dealing with entities, because they are able to participate as risk managers in reducing national systemic risk.

The difference between entities and assets is important. Although *entities* are firms, organizations, and operators, *assets* make up the components of NCFs. Therefore, although our objective is to identify SIEs, systemically important assets can be identified separately through NCF decomposition.

TABLE 2.1

The 55 National Critical Functions

Connect	Distribute	Manage	Supply
• Operate Core Network • Provide Cable Access Network Services • Provide Internet Based Content, Information, and Communication Services • Provide Internet Routing, Access, and Connection Services • Provide Positioning, Navigation, and Timing Services • Provide Radio Broadcast Access Network Services • Provide Satellite Access Network Services • Provide Wireless Access Network Services • Provide Wireline Access Network Services	• Distribute Electricity • Maintain Supply Chains • Transmit Electricity • Transport Cargo and Passengers by Air • Transport Cargo and Passengers by Rail • Transport Cargo and Passengers by Road • Transport Cargo and Passengers by Vessel • Transport Materials by Pipeline • Transport Passengers by Mass Transit	• Conduct Elections • Develop and Maintain Public Works and Services • Educate and Train • Enforce Law • Maintain Access to Medical Records • Manage Hazardous Materials • Manage Wastewater • Operate Government • Perform Cyber Incident Management Capabilities • Prepare for and Manage Emergencies • Preserve Constitutional Rights • Protect Sensitive Information • Provide and Maintain Infrastructure • Provide Capital Markets and Investment Activities • Provide Consumer and Commercial Banking Services • Provide Funding and Liquidity Services • Provide Identity Management and Associated Trust Support Services • Provide Insurance Services • Provide Medical Care • Provide Payment, Clearing, and Settlement Services • Provide Public Safety • Provide Wholesale Funding • Store Fuel and Maintain Reserves • Support Community Health	• Exploration and Extraction of Fuels • Fuel Refining and Processing Fuels • Generate Electricity • Manufacture Equipment • Produce and Provide Agricultural Products and Services • Produce and Provide Human and Animal Food Products and Services • Produce Chemicals • Provide Metals and Materials • Provide Housing • Provide Information Technology Products and Services • Provide Materiel and Operational Support to Defense • Research and Development • Supply Water

CHAPTER THREE

Systemically Important Entities: Identification and Prioritization

Recommendation 5.1 of the Cyberspace Solarium Commission report noted that CISA's responsibilities should include the identification of SICI (i.e., SIEs). Although the Solarium report did not say how SIEs could be identified, lessons from the designation of SIFIs, G-SIBs, and from the literature on systemic risk suggested potential directions. Additionally, although there is a clear need to identify SIEs, the potential number of SIEs could also range from the thousands to tens of thousands, suggesting a clear need to prioritize them.

In this chapter we describe a transparent, data-driven methodology for both identifying and prioritizing SIEs. Although cyber risk might have been the primary motivation for Solarium's recommendation, the concept of SIEs is broader than cyber risk, as shown in our approach to identification and prioritization. In addition, we describe the development of an interactive tool (using a Tableau visualization) that is designed to help compare the results of different priority criteria.

Identification of Systemically Important Entities

Our approach to identifying SIEs aims to be transparent, data driven, and fundamentally interoperable between data sets used today and data sets that could be used in the future. Broadly, we identify *SIEs* as entities that have the potential to disrupt one or many NCFs. To identify SIEs, we first identified thousands of entities that draw revenue from at least one NCF. Then we introduced a two-step process of (1) connecting NCFs to economic sectors and (2) connecting economic sectors to specific entities.

From National Critical Functions to Sectors

First, we map between NCFs and the (widely used) North American Industry Classification System (NAICS) of economic sectors. Although no definitive mapping exists between NCF and NAICS as of the writing of this report, we leveraged the most current mappings provided by the CISA OCE. Notably, several NCFs are grounded in noneconomic perspectives, which means that they cannot be mapped to NAICS sectors. For example, the NCF Preserve

Constitutional Rights does not have a clear economic analog. We do not classify the following ten NCFs into NAICS sectors:

1. Conduct Elections
2. Develop and Maintain Public Works and Services
3. Enforce Law
4. Operate Government
5. Prepare for and Manage Emergencies
6. Preserve Constitutional Rights
7. Protect Sensitive Information
8. Provide and Maintain Infrastructure
9. Provide Public Safety
10. Provide Material and Operational Support to Defense.

Although the tenth NCF on our list, Provide Material and Operational Support to Defense, does have an economic analog, it is dispersed across the majority of economic sectors. An entirely separate data-driven process using open-source data on defense contracts can be used to identify all firms providing material and operational support to DoD.

Next, we adopt a three-step approach for mapping the remaining NCFs to one or more NAICS sectors. Importantly, the NAICS taxonomy is a hierarchical structure so there is an inherent nesting structure within the taxonomy, and not all the NCFs are within the same level within the NAICS taxonomy. For example, Produce and Provide Agricultural Products and Services corresponds to a two-digit NAICS sector, Transport Materials by Pipeline corresponds to a three-digit NAICS sector, and Manage Wastewater corresponds to a five-digit NAICS sector. Consequently, our method looks across all possible levels of NAICS codes.

We then use the NCF to USA General Equilibrium (USAGE) model mapping of NCFs to the sectors and the NAICS to USAGE mapping provided by OCE. However, given that the NCF to USAGE mapping is done at a very high level of aggregation because of USAGE sectors, many NCFs map to a single USAGE sector, which roughly corresponds to three-digit NAICS sectors (e.g., all telecommunications and broadcasting NCFs map to a single USAGE sector). Thus, we take additional steps to map at lower levels of aggregation.

Next, because of the nesting structure of the NAICS taxonomy, we use the complete set of six-digit NAICS codes that correspond to the USAGE sector that the NCF was mapped to as a starting point and eliminate sectors that do not correspond to the NCF. This provides an initial mapping from NCFs to NAICS sectors that is consistent with an aggregation to the USAGE sectoring and is consistent with the OCE mappings from NCF to NAICS.

We then search for other NAICS codes that could be included in the NCF that might not neatly fit within the USAGE sectoring because it might make up a small subsector of the broad USAGE sector. This applies to only very few NCFs as there is generally a clear mapping from NCF to NAICS for those NCFs that have economic analogs.

Finally, we independently validate the mapping by enlisting an outside economist to review and discuss the NAICS codes that were included and excluded from our mappings. This allows the mapping to be validated by someone outside the project who does not have a vested interest in the project. The final mapping is provided in Table 3.1.

TABLE 3.1
National Critical Function to North American Industry Classification System Mapping

NCF	NCF Description	NAICS
Operate Core Network	Maintain and operate communications backbone infrastructure for voice video and data transmission that connects to users through broadcasting cable satellite wireless and wireline access networks	517: Telecommunications 518: Data Processing, Hosting and Related Services 5415: Computer System Design and Related Services 8112: Electronic and Precision Equipment Repair and Maintenance
Provide Cable Access Network Services	Provide access to communications backbone infrastructure through fiber and coaxial cable network supplying analog and digital video programming services digital telephone service and high-speed broadband services	517311: Wired Telecommunications Carriers
Provide Internet Based Content, Information, and Communication Services	Produce and provide technologies services and infrastructure that deliver key content information and communications capabilities via the internet	51913: Internet Broadcasting and Web Search Portals
Provide Internet Routing, Access, and Connection Services	Provide and operate exchange and routing infrastructure points of presence peering points local access services and capabilities that enable end users to send and receive information via the internet	541511: Custom Computer Programming Services
Provide Positioning, Navigation, and Timing Services	Operate and maintain public and private capabilities which enable users to determine location orientation and time	517919: All Other Telecommunications
Provide Radio Broadcast Access Network Services	Operate over-the-air radio and television (TV) stations (operating at medium very high and ultra-high frequencies) that offer analog and digital audio and video programming services and data service	515: Broadcasting Except Internet
Provide Satellite Access Network Services	Provide access to core communications network via a combination of terrestrial antenna stations and platforms orbiting Earth to relay voice video or data signals	5174: Satellite Telecommunications

Table 3.1—Continued

NCF	NCF Description	NAICS
Provide Wireless Access Network Services	Provide access to core communications network via electromagnetic wave-based technologies, including cellular phones, wireless hot spots (Wi-Fi), personal communication services, high-frequency radio unlicensed wireless, and other commercial and private radio services	517312: Wireless Telecommunications Carriers
Provide Wireline Access Network Services	Operate circuit-and packet-switched networks via copper fiber and coaxial transport media, including private enterprise data and telephony networks and the public switched telephone network (PSTN)	517311: Wired Telecommunications Carriers
Distribute Electricity	Maintain and operate medium- to low-voltage system to reliably supply consumer demand for electricity from the bulk electric power network	221122: Electric Power Distribution
Maintain Supply Chains	Manage and sustain the networks of assets systems and relationships that enable the movement of goods and services from producers to consumers	48–49: Transportation and Warehousing
Transmit Electricity	Maintain and operate high-voltage (>100kV) bulk electric system to reliably supply distribution network demand for electricity from generation resources	221121: Electric Power Transmission
Transport Cargo and Passengers by Air	Provide and operate aviation systems assets and facilities to enable a system to securely and safely convey goods and people from place to place by air	481: Air Transportation
Transport Cargo and Passengers by Rail	Provide and operate freight and passenger railroad systems including conveyances infrastructure and management systems to enable a system to securely and safely convey goods and people from place to place by rail	482: Rail Transportation

Table 3.1—Continued

NCF	NCF Description	NAICS
Transport Cargo and Passengers by Road	Provide and operate roadway systems assets and facilities including commercial motor carriers and associated facilities motor coaches buses and associated systems assets and facilities to enable a system to securely and safely convey goods and people from place to place by highway	484: Truck Transportation
Transport Cargo and Passengers by Vessel	Provide and operate maritime systems, assets, and facilities to enable a system to securely and safely convey goods and people from place to place by the Maritime Transportation System	483: Water Transportation
Transport Materials by Pipeline	Provide and operate systems assets and facilities to enable a system to securely and safely convey materials from place to place by pipelines	486: Pipeline Transportation
Transport Passengers by Mass Transit	Provide and operate systems assets and facilities to enable a system to securely and safely convey people from place to place by roads or on fixed guideways within a specified geographic area, including transit buses, trolleybuses, monorails, heavy rail (subway), light rail, passenger rail, commuter rail	485: Transit and Ground Passenger Transportation
Educate and Train	Provide education and workforce training including Pre-K-12, community college, university, and graduate education technical schools with apprenticeships, nonformal education, and on-the-job training	61: Educational Services
Maintain Access to Medical Records	Maintain, use, and share actionable data (including personally identifiable information and personal health information, such as care history) effectively, appropriately, bidirectionally, and in a timely fashion for patient care billing and operational and clinical research	621: Ambulatory Health Services 622: Hospitals 623: Nursing and Residential Care Facilities
Manage Hazardous Materials	Safely identify, monitor, handle, store, transport, use, and dispose of hazardous materials (including chemical biological radioactive)	562211: Hazardous Waste Treatment and Disposal

Table 3.1—Continued

NCF	NCF Description	NAICS
Manage Wastewater	Collect and treat industrial and residential wastewater to meet applicable public health and environmental standards prior to discharge into a receiving body	221320: Sewage Treatment Facilities
Perform Cyber Incident Management Capabilities	Provide security systems and services that protect critical business assets and functions, including preventive guidance simulation testing and warning capabilities; operate operations response centers and teams; integrate and share information; coordinate and provide response recovery and reconstitution services	541519: Other Computer Related Services 5616: Investigation and Security Services
Provide Capital Markets and Investment Activities	Issue and trade securities, including debt securities (such as bonds), equities (such as stocks), and derivatives (such as options and futures); provide advisory services and related services, such as prime brokerage; maintain and operate organized markets and over-the-counter mechanisms for these instruments	523: Securities, Commodity Contracts, and Other Financial Investment and Related Activities 525: Funds, Trusts, and Other Financial Vehicles
Provide Consumer and Commercial Banking Services	Accept and maintain deposit accounts (e.g., checking and savings accounts) and close substitutes (e.g., short-term retail notes) from nonfinancial intermediaries	522: Credit Intermediation and Related Activities
Provide Funding and Liquidity Services	Provide funding to nonfinancial counterparties, such as corporate or retail customers, including individual consumers	522: Credit Intermediation and Related Activities
Provide Identity Management and Associated Trust Support Services	Produce and provide technologies, services, and infrastructure to ensure the ability to identify, authenticate and authorize entities and ensure confidentiality, integrity, and availability of devices' service data	5415: Computer System Design and Related Services
Provide Insurance Services	Operate systems and markets to transfer financial risks among parties through contractual relationships including products for individuals, corporations, and public-sector entities	524: Insurance Carriers and Related Activities

Table 3.1—Continued

NCF	NCF Description	NAICS
Provide Medical Care	Ensure the provision of health care services	621: Ambulatory Health Services 622: Hospitals 623: Nursing and Residential Care Facilities
Provide Payment, Clearing, and Settlement Services	Carry out processes required for the exchange of assets, including payment (transfer of funds between or among participants), clearing (transmitting, reconciling, and confirming transactions prior to settlement), and settlement (transfer of ownership and payments)	52232: Financial Transactions Processing, Reserve and Clearing House Activities
Provide Wholesale Funding	Maintain processes for lending and borrowing among financial services sector parties	52211: Commercial Banking
Store Fuel and Maintain Reserves	Store energetic materials (including fossil and nuclear fuels) to reliably meet operational and strategic demands	42471: Petroleum Bulk Stations and Terminals 48621: Pipeline Transportation of Natural Gas
Support Community Health	Conduct epidemiologic surveillance, environmental health, migrant and shelter operations, food establishment inspections, and other community-based public health activities	6242: Emergency and Other Relief Services 92312: Administration of Public Health Programs
Exploration and Extraction of Fuels	Identify resources and collect energetic materials (including fossil fuels, nuclear materials, and others)	2111: Oil and Gas Extraction 212291: Uranium Ore Mining 325180: Other Basic Nonorganic Chemical Manufacturing
Fuel Refining and Processing Fuels	Transform raw energetic materials into consumer fuels (e.g., crude cracking gas separation and uranium enrichment)	324110: Petroleum Refiners 325180: Other Basic Nonorganic Chemical Manufacturing
Generate Electricity	Produce electricity from a variety of primary energy sources (including fossil fuels, nuclear materials, and renewables) to reliably meet demand	22111: Electric Power Generation
Manufacture Equipment	Fabricate and assemble components to produce tangible property	33: Manufacturing
Produce and Provide Agricultural Products and Services	Grow and harvest plant and animal commodities (including crops, livestock, dairy, aquaculture, and timber) and produce inputs required to support agricultural production (such as fertilizers, pesticides, animal food, crop seeds, and veterinary services)	11: Agriculture, Forestry, Fishing, and Hunting

Table 3.1—Continued

NCF	NCF Description	NAICS
Produce and Provide Human and Animal Food Products and Services	Produce food products from raw agricultural commodities and provide to final consumers (including processing, packaging, production, product storage, and retail and food service)	311: Food Processing
Produce Chemicals	Manufacture basic chemicals from raw organic and inorganic materials and manufacture intermediate and final products from basic chemicals	325: Chemical Manufacturing
Provide Metals and Materials	Manufacture iron steel and ferroalloy products, alumina and aluminum products, nonferrous metals, and other materials as primary components for other industries	3321: Forging and Stamping 3322: Cutlery and Hand Tool Manufacturing 3323: Architectural and Structural Metals Manufacturing
Provide Housing	Construct and/or provide safe and secure permanent or temporary shelter for people (includes physical construction and emergency sheltering)	2361: Residential Building Construction 6242: Emergency and Other Relief Services
Provide Information Technology Products and Services	Design, develop, and distribute hardware and software products and services (including security and support services) necessary to maintain or reconstitute networks and associated services	5112: Software Publishers 334111: Electronic Computer Manufacturing 238210: Electrical Contractors and Other Wiring Installation Contractors
Research and Development	Conduct basic research; innovate, test, and introduce new products and services; or improve existing products and services	5417: Scientific Research and Development Services
Supply Water	Maintain availability of water (raw and treated)	22131: Water Supply and Irrigation Systems

SOURCE: CISA, *National Critical Functions: Status Update to the Critical Infrastructure Community*, Washington, D.C.: U.S. Department of Homeland Security, July 2020a.

There are a few things to note about this mapping. First, multiple NCFs map to the same NAICS codes or subsets of the NAICS codes. For example, it is difficult to disentangle the provision of medical care from medical records. Similarly, payment clearing and settlement services is a subcomponent of the banking industry. Furthermore, the inability to disentangle some NCFs from each other and the reality that some NCFs are subcomponents of other NCFs makes any effort at comparing analyses across NCFs challenging. That is, the simple fact that a sector might appear in multiple NCFs does not make that sector inherently more important than another.

From Sectors to Entities

To identify which entities are associated with each sector and to NCFs through the mapping described above, we leveraged the commercial business data set, FactSet. **Specifically, we used data from FactSet and its taxonomy, which is called the Revere Business Industry Classification System (RBICS), to tie entities to sectors based on their revenues.** FactSet provided RAND with a mapping of six-digit NAICS codes to the corresponding level 6 RBICS sectors, two taxonomies that operate at roughly the same scale of the economy. Thus, we take advantage of the nesting nature of the NAICS codes and RBICS codes to create the mapping from NCF to RBICS using the NCF to NAICS as the backbone of the mapping. Although we have done this using the FactSet data, the approach is fundamentally interoperable using other data services, such as Bloomberg, S&P Capital IQ, or Dun & Bradstreet.

We use FactSet as our underlying data set for three reasons that will become more apparent in the next section. First, FactSet data provide all RBICS sectors in which an entity derives revenue. In many cases, only a single NAICS code will be provided for a firm in other data sets. Second, FactSet data provide estimates of the revenue associated with each of the RBICS sectors. Third, FactSet data provide a gold standard set of business relationships (i.e., supply chain connections), relationships that are based on official company statements on investor disclosures, press releases, transcripts, presentations, and other official records. FactSet data include not only the relationships that an entity discloses but the relationships that are disclosed by partner entities. That is, if firm A discloses that they are a supplier to firm B, we automatically know that firm B is a customer of firm A.

This approach provides a list of thousands of entities with the potential to become SIEs.

Prioritization of Systemically Important Entities

Although the identification step provides a list of thousands of entities with the potential to become SIEs, the prioritization step provides a methodology for sorting entities by measures of their systemic importance. Specifically, we drew inspiration from the supporting methodology for identifying SIFIs and G-SIBs based, in part, on their size and interconnectedness.

To estimate an entity's size, we estimate which entities appear to be important economically within the NCF based on their sector-specific revenues. That is, we leverage the mapping created in the previous section which identifies entities in association with RBICS sectors with RBICS-provided revenue by sector to calculate an entity's corresponding revenue by NCF.

Next, to understand an entity's interconnectedness, we estimated a given entity's centrality in economic networks. This approach leverages the work of Welburn, Strong, et al. (2020) in calculating the centrality of firms in interfirm networks. Here, we exploit data provided by FactSet Revere, a large data feed of customer and supplier relationships. From this, we construct the corresponding interfirm network, represented by an adjacency matrix, associated with all of the customer-supplier relationships.

In constructing the interfirm network, let A be this adjacency matrix such that $a_{ij} = 1$ if entity j is a customer of entity i. Measures of centrality are used to rank nodes in a network based on their importance to other nodes within that network. Perhaps the simplest measure of centrality is given by its degree: a count of the number of connections each node has directly to any other given node. An entity with three customers and two suppliers in the network, for example, has a degree of 5. The degree of each node is calculated simply as

$$AI = x$$

where x is a vector of degrees, x_i for each entity and all entities can be ranked according to the highest degree or its degree centrality, and I is an identity column vector. Although simple in explanation, degree centrality does not take into consideration the full structure of the interfirm network.

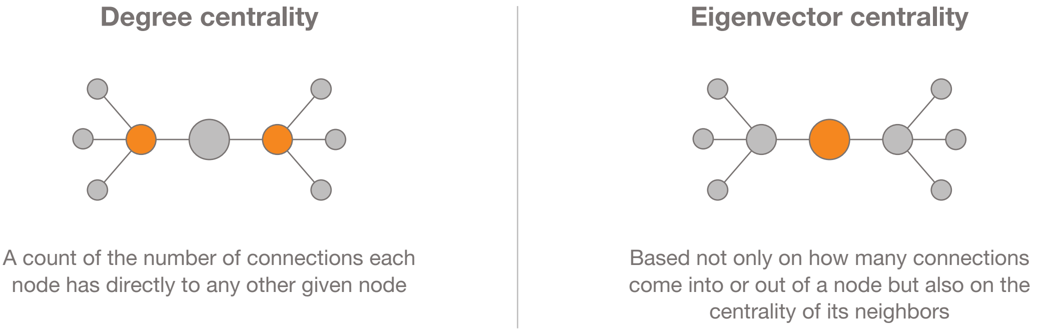
Consequently, we use another widely used measure of centrality—eigenvector centrality—to measure the importance of an entity within the broader interfirm network based on its interconnectedness.[1] Figure 3.1 visualizes the difference between degree and eigenvector centrality, where the most-central nodes under each method are shown in orange. Unlike degree centrality, eigenvector centrality is based not only on how many connections come into or out of a node but also the centrality of its neighbors. Eigenvector centrality x is defined as the vector that solves

$$Ax = \lambda x, \text{ with } x = [x_1 \ldots x_n]$$

where λ is the largest eigenvalue associated with the eigenvector x. Each of the x_i's is the eigenvector centrality of node i.

There are other possible measures of centrality beyond degree and eigenvector. Two other common measures of centrality and closeness prioritize nodes by their paths and are focused

FIGURE 3.1
Network Centrality



A count of the number of connections each node has directly to any other given node

Based not only on how many connections come into or out of a node but also on the centrality of its neighbors

[1] Britta Ruhnau, "Eigenvector-Centrality—A Node-Centrality?" *Social Networks*, Vol. 22, No. 4, 2000.

on the shortest paths between pairs of nodes. Furthermore, the PageRank algorithm, pioneered by Google's Larry Page to rank search results based on links across pages, is a famous variant of eigenvector centrality.

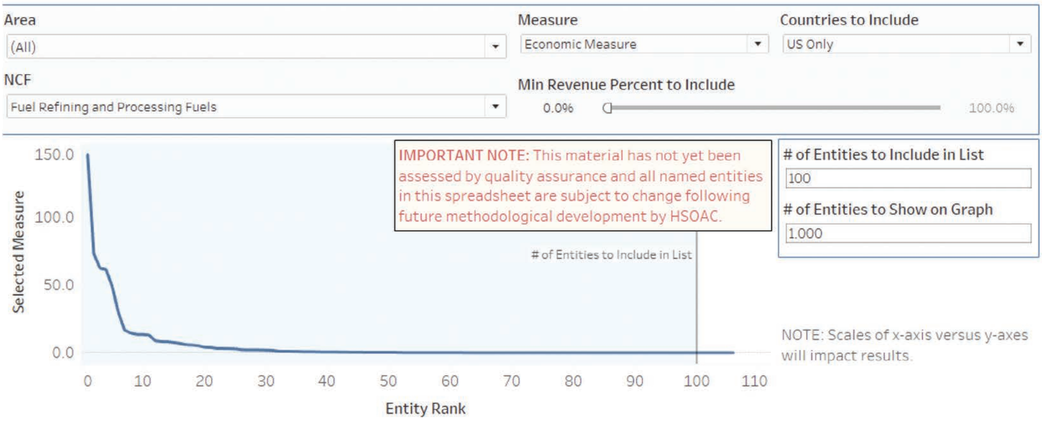
Given its value for solving our specific problem of estimating interconnectedness, we use eigenvector centrality as our measure of network centrality.

We recommend prioritizing potential SIEs based off the two measures of size and interconnectedness. Naturally, one could be interested in a single combined metric. We provide a basic combined metric by dividing both measures by their maximum values. Because some foreign entities might enter the list, we allow for two different maximum NCF revenues, one based on the largest U.S. firm and the other based off the largest global firm that falls in the NCF. After normalization, an equal weighting of revenues and network centrality is likely the most logical. However, we provide analysts and decisionmakers with the ability to adjust weights independently in an interactive tool. For example, for the NCF Distribute Electricity, the number of customers as proxied by revenue might be more important than network centrality because it provides a better measure of the potential impact of a disruption. Conversely, for the NCF Provide Internet Routing Services, it might be more informative to know the network centrality because the size of the relationship does not necessarily matter but the number of connections does (because any disruption would be disabling to the service, no matter the scale).

Furthermore, in the interactive tool (SIAM), we provide illustrations of how the concentration of entities within an NCF, provided by the distribution of their measures of size and interconnectedness, can provide insight into how many entities are systemically important within each NCF. Figure 3.2 depicts the distribution of entity revenue across the NCF Fuel Refining and Processing. Moving from left to right, the curve sharply decreases at first, before gradually tapering off and eventually converging to zero revenue. Specifically, the curve begins to taper off and flatten at a little less than ten entities, a point that we call the "knee-in-the-curve." This suggests that the number of entities that should be considered as

FIGURE 3.2

Knee-in-the-Curve Graph for Fuel Refining and Processing



Area: (All)

Measure: Economic Measure

Countries to Include: US Only

NCF: Fuel Refining and Processing Fuels

Min Revenue Percent to Include: 0.0% – 100.0%

IMPORTANT NOTE: This material has not yet been assessed by quality assurance and all named entities in this spreadsheet are subject to change following future methodological development by HSOAC.

# of Entities to Include in List: 100

# of Entities to Show on Graph: 1,000

NOTE: Scales of x-axis versus y-axes will impact results.

an SIE (using size as an indicator of systemic importance) is likely less than 10 for this specific NCF. If no obvious knee-in-the-curve exists, the number of potentially relevant entities will be large, and greater subject-matter expertise should be sought.

Additionally, a measure of economic centrality could produce a single combined, global measure of both size and interconnectedness. This metric could leverage not only the existence of connections between entities, but their weight based on the dollar value of flows between them. Although no complete data sets exist on flows between entities, future work can follow the methodology of Welburn, Strong, et al., (2020) to estimate them and, subsequently, to estimate an entity's economic centrality.[2]

Limitations

Our current approach is focused on economic importance. Therefore, we are unable to provide insight into SIEs for such NCFs as Enforce Constitutional Rights. Because there is no information that can be gleaned from this analysis, we have removed them from consideration. Additional perspectives could be incorporated into successive iterations of this analysis. The general approach of using network analysis of the global system together with an NCF metric views the NCFs as a system in themselves while recognizing that the NCF is part of a larger system-of-systems (either from the perspective of an NCF or alternative goals, such as health and well-being, homeland security, national security, or any other perspective that would need to be incorporated into systemic importance). In some cases, it might make sense to consider only one of these dimensions (either because the NCF is relatively self-contained as a system and not need a global interconnectedness measure, or because there are no clear analogs to the NCF revenue for the specific NCF).

Although the underlying data set includes several types of entities, including privately held firms, it is quite skewed toward publicly traded firms. This is largely an artifact of stronger disclosure requirements for publicly traded firms in the United States, where private firms are not compelled to provide insight into financials or interconnections. That said, our analysis does provide some insight into other types of firms. Table 3.2 depicts the representation of 19 entity types included in our data set of potential SIEs. Of them, a considerable majority (76 percent) are public companies, 17 percent are subsidiaries,[3] 4 percent are privately held companies (a number that often captures larger privately held companies), 3 percent are holding companies, and notably small numbers of entities are dispersed across the other 15 entity types.

[2] Jonathan W. Welburn, Aaron Strong, Florentine Eloundou Nekoul, Justin Grana, Krystyna Marcinek, Osonde A. Osoba, Nirabh Koirala, and Claude Messan Setodji, *Systemic Risk: It's Not Just in the Financial Sector*, Santa Monica, Calif.: RAND Corporation, RB-10112-RC, 2020.

[3] A subsidiary is an entity that is wholly owned by another entity, which is often referred to as the parent company.

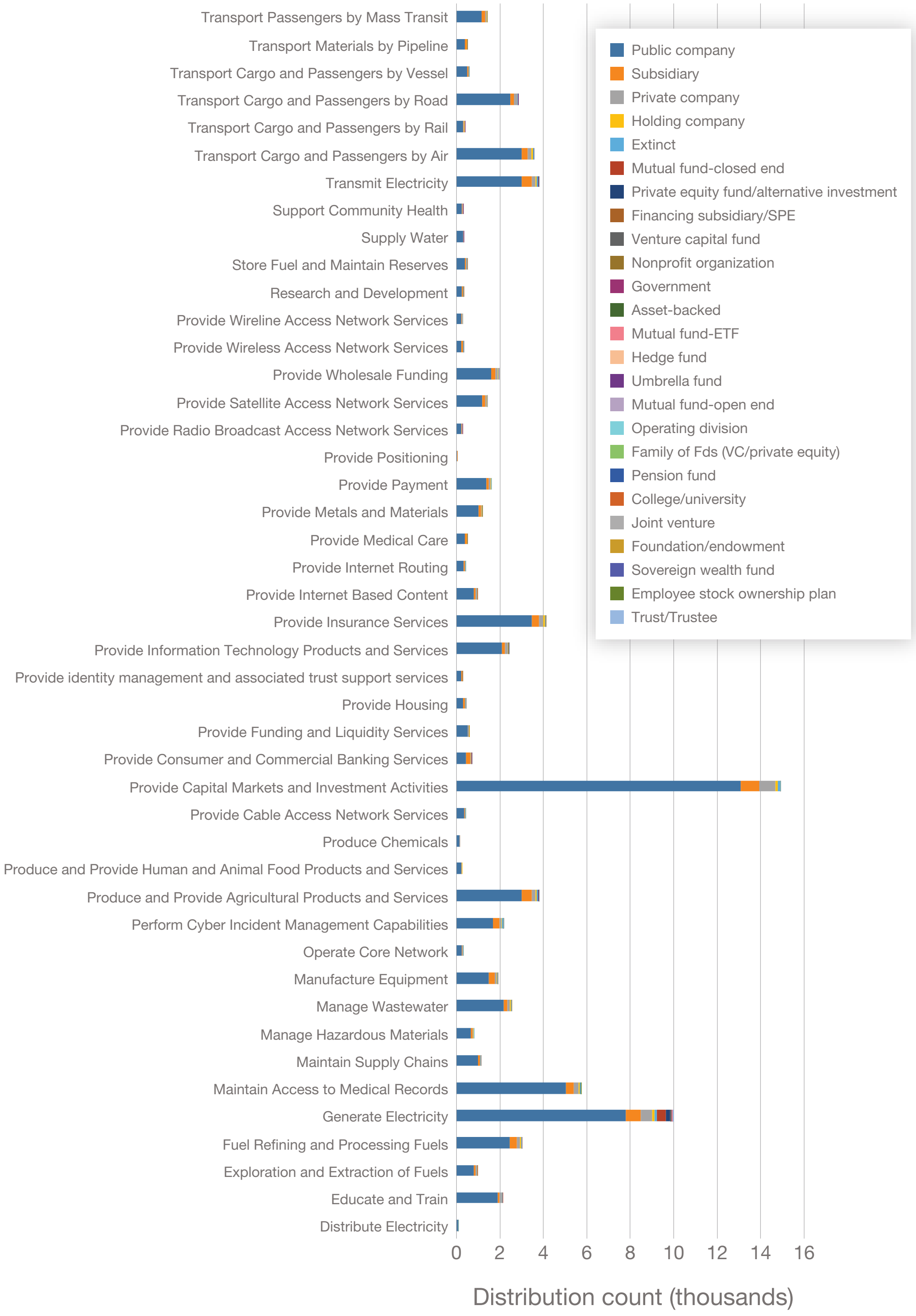
TABLE 3.2
Share of Entities, by Entity Type

Entity Type		Percentage
Public Company	A publicly traded company	75.51
Subsidiary	An entity whole owned by another entity	16.76
Private Company	A privately held company that is not traded publicly	3.93
Holding Company	A company the holds shares of other companies	2.51
Private Equity Fund/ Alternate Investment	A private company that uses investor money to invest in or acquire ownership in a company	0.44
Mutual Fund/Closed End	A mutual fund with a fixed number of shares	0.38
Financing Subsidiary/ Special Purpose Entity	A separate legal entity created solely to carry out financial transactions	0.13
Nonprofit Organization	An entity organized around activities other than making a profit and none of the organization's income is distributed to members, directors or officers	0.10
Government	A sovereign entity working on behalf of its citizens	0.07
Venture Capital Fund	Pooled investment funds that seek private equity stakes in start-ups and small-to-medium-sized operations with strong growth potential	0.04
College/University	An entity of higher education beyond that of high school	0.03
Joint Venture	An entity backed by two or more entities that maintain distinct identities	0.03
Asset-Backed	Asset-backed entities that are created when a company's loans are packaged as a portfolio to sell to investors	0.01
Hedge Fund	A limited partnership of investors	0.01
Operating Division	A separate division of a company that is operated independently but under the parent company	0.01
Foundation/Endowment	An independent entity engaged in charitable purposes	0.01
Mutual Fund-Open End	A mutual fund that does not have a fixed number of shares and can be bought and sold on demand	0.01
Mutual Fund-ETF	A mutual fund that usually tracks a particular index and can be traded on a stock exchange	0.01
Umbrella Fund	An investment fund that has multiple sub-funds	0.01

Figure 3.3 depicts the distribution of each entity type across NCFs. Unsurprisingly, finance-specific entities, such as hedge funds, are concentrated in finance-specific NCFs, such as Provide Consumer and Commercial Banking Services and Provide Housing. Although

FIGURE 3.3

Distribution of Entity Types Across Natonal Critical Functions



firms (i.e., publicly and privately traded entities, holding companies) are more concentrated in certain NCFs, such as Operate Core Network, this could either indicate a difference in the composition of NCFs or a difference in transparency and data collection by NCF. Additionally, Figure 3.3 highlights the meaningful variance of potential SIEs that have been identified across NCFs, with the most in Operate Core Network and the least in Transmit Electricity.

Finally, our approach is focused on a national perspective, which might not be the appropriate scale for some of the NCFs, because the delivery system is local, not national. For those NCFs that have a regional delivery system, it might be necessary to further disaggregate the NCF into regional components (e.g., the Federal Energy Regulatory Commission's regions for the Generate, Transmit, and Distribute Electricity NCFs). Similarly, in terms of systemic importance, Supply Water is probably best viewed at the local or regional level rather than the national level.

Overcoming Limitations from an Equity Perspective

Alternative approaches to systemic importance that are not purely economic might help overcome the limitations discussed above. These alternative approaches can help identify entities that are essential for people's well-being or for interests, such as national security even if those entities do not have large aggregate economic values. We briefly lay out an alternative approach that identifies systemic importance from an equity perspective. Notably, this perspective is predominately focused on supply side issues that follow from a significant disruption to entities of systemic importance. This discussion will underscore the importance of this perspective and chart a path forward for detailed equity analysis that can be combined with the economic approach and other perspectives to offer a more complete view of systemic importance.

The Biden administration has committed to developing policies and programs that address historic and present-day inequities by prioritizing underserved communities. This commitment is reflected in presidential directives, senior-level statements, and executive branch department strategy documents, among other places. For instance, Executive Order (EO) 13985 "Advancing Racial Equity and Support for Underserved Communities," which was released on President Biden's first day in office, calls on federal departments and agencies to address inequities and promotes a whole-of-government equity agenda.[4] The administration has sought to implement this commitment to equity in policy and such programs as the Justice40 initiative, which requires that 40 percent of all investments in climate change, sustainable infrastructure, and clean water accrue to disadvantaged communities.[5] In addition,

[4] EO 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government," January 20, 2021.

[5] Shalanda Young, Brenda Mallory, and Gina McCarthy, "The Path to Achieving Justice40," *The White House Briefing Room Blog*, July 20, 2021.

agency heads, such as DHS Secretary Mayorkas and others, have affirmed the importance of equity in the context of significant events (such as disasters) where there is a long-standing challenge of underserved communities facing disparate impacts and barriers in obtaining assistance.[6] As Mayorkas states, "Equity is a cornerstone of our homeland security mission and in all of our work we must reach minority communities, the disadvantaged, and the otherwise disenfranchised."[7]

The economic approach to systemic importance discussed above does not take into account, much less prioritize, communities that are underserved and marginalized. In general, these communities are invisible or overlooked in the economic and network centrality approaches to systemic importance because they have fewer assets and less wealth. As a result, they will not constitute a significant portion of aggregate economic metrics, while those with greater resources will be disproportionally represented. Unfortunately, risks stemming from cyber incidents or other disruptions to underserved communities might be particularly severe, especially because these communities have fewer resources with which to withstand shocks and recover. Moreover, because a dollar loss to someone who is poor is worse than a dollar loss to someone who is rich, the losses from disruption will generally reflect a greater share of disadvantaged persons' well-being than of those communities that enjoy relative advantages. When disruptions occur, communities that are already disadvantaged will be even more negatively impacted, especially if they are not prioritized in analysis and policy.

Given these issues, there have been attempts to understand the impacts of disasters and other disruptions from the standpoint of how they affect individuals' well-being, rather than through purely economic metrics.[8] This analysis underscores that it is not just the total amount of loss from a disruption that matters, but how those losses are distributed across the population and the ways the loss impacts actual communities, including the safety net of resources and government assistance that helps them be resilient.

The entities that serve marginalized communities might not be as economically important in the aggregate and might not have significant reach through established business networks. Or these entities might be primarily important in specific regions or localities with large percentages of underserved populations while not having a major role across regions or at a national scale. The systemic importance of these entities might be related not to their aggregate economic relationships, but to who they serve.

[6] See Christopher Flavelle, "Why Does Disaster Aid Often Favor White People?" *New York Times*, June 7, 2021.

[7] DHS, "DHS Announces Changes to Individual Assistance Policies to Advance Equity for Disaster Survivors," press release, September 2, 2021.

[8] Stephane Hallegatte, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg, *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters*, Climate Change and Development, Washington, D.C.: World Bank, 2017.

In addition, specific NCFs, such as Prepare for and Manage Emergencies or Develop and Maintain Public Works and Services that are not included in the economic analysis above, might be particularly important for underserved communities, and thus there is a need to identify the SIEs that relate to those NCFs.

Beyond the ethical case for prioritizing underserved communities, equity connects individuals to opportunities, resources, and networks that have implications for the overall U.S. economy. As stated earlier, systems are essentially a network of different entities. Similarly, the U.S. economy is a system composed of a network of individuals who contribute to the production and consumption of goods and services. Considering that individuals can vary on such demographics as race, ethnicity, sex, sexual orientation, or disability status, examining the extent to which subgroups of the population are underserved in terms of opportunities and resources has direct and indirect implications for the aggregate U.S. economy. A 2020 report by Citi found that not addressing the racial gap between Black and White Americans has cost the U.S. economy up to $16 trillion over the past 20 years.[9] Despite policies enacted by the federal government that prohibit discrimination, the legacy of inequities in various sectors across the United States, including health, education, housing, and the labor market, continue to impact quality of life for underserved groups.[10]

For instance, examining trends in the Black-White wage gap shows that a significant portion of the gap is attributable to unexplained factors unrelated to occupational choice, educational attainment, and age, with unexplained factors accounting for a larger share of the wage gap over time.[11] Closing the Black-White wage gap could result in an additional $2.7 trillion in income in the economy for consumption or investment.[12] Additionally, improving access to housing credits and incorporating fair and equitable lending practices could add $28 billion in homeownership sales and expenditures and $13 trillion in business revenue that could result in the creation of 6.1 million jobs per year.[13] Overall, equity for underserved communities has implications beyond moral implications related to social justice and fairness. Equity for these communities has broad implications for the overall U.S. economy, resulting in an additional $5 trillion in gross domestic product over the next five years.[14]

How can analysis of systemic importance be supplemented with an equity perspective? Equity is a complex concept with a multitude of meanings and is often defined in relation to

[9] Dana M. Peterson, and Catherine L. Mann, *Closing the Racial Inequality Gaps: The Economic Cost of Black Inequality in the U.S.*, Citi GPS: Global Perspectives and Solutions, September 2020.

[10] Ani Turner, "The Business Case for Racial Equity," *National Civic Review*, Vol. 105, No. 1, 2016.

[11] Mary C. Daly, Bart Hobijn, and Joseph H. Pedtke, "Disappointing Facts About the Black-White Wage Gap," *FRBSF Economic Letter*, Vol. 26, 2017; Eleni Karageorge, "The Unexplainable, Growing Black-White Wage Gap," *Monthly Labor Review*, Vol. 140, November 2017.

[12] Peterson and Mann, 2020.

[13] Peterson and Mann, 2020.

[14] Peterson and Mann, 2020.

other complex normative concepts that are hard to define (such as "justice" or "fairness").[15] Equity is regularly distinguished from pure equality of resources or outcomes; instead of treating everyone exactly the same or giving each person the same resources, equity requires understanding and addressing the relative positionality and set of benefits and burdens individuals face within shared institutional structures. The goal is to achieve a just set of social institutions in which any resulting advantages and disadvantages are based on compelling moral foundations (e.g., free choice or need) rather than on unjust historical and present-day conditions (e.g., racist policies, such as red-lining) or patriarchal policies that have hindered women. However, even with this objective in mind, our everyday intuitions about what equity requires are sometimes in tension, both across and within individuals. Although some arguments might suggest a wholesale change in policy to advance equity, institutional changes need to be considered alongside other values that might be relevant, including individual rights and liberties.

The U.S. government and other actors seeking to achieve equity typically focus on underserved, marginalized, or disadvantaged communities because these are the communities that have not received their share of the benefits of social cooperation. However, despite the overarching commitment to equity by the U.S. government, departments and agencies have not been clear about the specific actions or policies that advance equity, the communities that should be prioritized, or the metrics or standards used to assess whether the actions or policies are effective.[16]

Fortunately, there are valuable frameworks that help unpack and make the complex, multifaceted concept of equity more concrete, and we adopt a distinction between the following three types of equity that have been used in related contexts:[17]

- *procedural equity:* Who is included or left out of important policy decisions?
- *contextual equity:* What political, cultural, economic, historical, or other factors exclude or marginalize individuals?
- *distributional equity:* What is the distribution of the benefits and burdens across individuals and communities?

[15] "The term 'equity' means the consistent and systematic fair, just, and impartial treatment of all individuals" (see Executive Order 13985, 2021).

[16] Noreen Clancy, Melissa L. Finucane, Jordan R. Fischbach, David G. Groves, Debra Knopman, Karishma V. Patel, and Lloyd Dixon, *The Building Resilient Infrastructure and Communities Mitigation Grant Program: Incorporating Hazard Risk and Social Equity into Decisionmaking Processes*, Homeland Security Operational Analysis Center operated by the RAND Corporation, RR-A1258-1, 2022.

[17] Melanie McDermott, Sango Mahanty, and Kate Schreckenberg, "Examining Equity: A Multidimensional Framework for Assessing Equity in Payments for Ecosystem Services," *Environmental Science and Policy*, Vol. 33, 2013; Melissa L. Finucane, Linnea Warren May, and Joan Chang, *A Scoping Literature Review on Indicators and Metrics for Assessing Racial Equity in Disaster Preparation, Response, and Recovery*, Santa Monica, Calif.: RAND Corporation, RR-A1083-1, 2021.

With this framework in mind, an attempt to identify SIEs from the equity perspective would need to identify the entities most essential for the communities that have been excluded from key decisionmaking, that face significant economic, political, and social barriers, and that have not received their fair share of the benefits from social cooperation. The analysis will need to be specific about the relevant underserved communities, the barriers they face, the allocation of benefits and burdens, and the entities that are particularly important for improving their well-being, all in relation to specific NCFs.

Example: How Systemic Importance Can Be Analyzed from an Equity Perspective

To begin to move forward in the analysis, we will discuss a specific example related to water which touches on NCFs that are included in the economic analysis (e.g., Supply Water), and also NCFs that are not (Develop and Maintain Public Works and Services). This example is intended to demonstrate the importance of the equity analysis, how it can be conducted, and how intersecting factors compound to create barriers to underserved communities within specific NCFs and also lead to systemic effects across NCFs.

Clean water has been a central focus of conversations about infrastructure access in disadvantaged communities. Although access to clean water and functioning sanitation systems is better in the United States than in almost any other country in the world, the U.S. Water Alliance found that more than 2 million Americans still lack access to clean water and functioning sanitation systems. Specific challenges include contaminated water, limited or no indoor plumbing, and limited or no access to wastewater systems.

Households that lack access to adequate clean water and sanitation are more likely to be in rural communities than those with adequate access to plumbing.[18] In addition, the U.S. Water Alliance (2019) found that people who lack access to adequate sanitation are more likely to be racial minorities.[19] Native Americans are most likely to lack access, followed by African Americans and Latinos.

The distributional inequity surrounding water can lead to several problems, including poor health, and can also perpetuate racial and economic inequality by diminishing the social mobility prospects of children and adults. For example, children who become infected with certain diseases (such as hookworm) are likelier than those who do not to experience developmental delays. These developmental delays can have a negative impact on educational and occupational attainment, which in turn limits opportunities to move up the socioeconomic ladder. This can add to a host of existing barriers for children who are living in low-income areas, making it particularly challenging to escape persistent poverty. Adults who suffer from diseases related to lack of access to clean water can also suffer short-term and

[18] J. Tom Mueller and Stephen Gasteyer, "The Widespread and Unjust Drinking Water and Clean Water Crisis in the United States," *Nature Communications*, Vol. 12, No. 1, 2021.

[19] U.S. Water Alliance, "Closing the Water Access Gap in the United States: A National Action Plan," 2019.

long-term health problems. Persistent health problems can make it difficult for these individuals to complete daily tasks, fully participate in the labor force, and pursue opportunities for educational and occupational growth. Thus, their mobility prospects can be severely limited. In these ways, the inequities associated with water-related NCFs themselves have systemic effects across a variety of other NCFs.

There are almost 70,000 water systems in the United States.[20] A 2014 survey conducted by the Alabama Center for Rural Enterprise finds that between 40 percent and 90 percent of residents in Lowndes County, Alabama, lack access to a fully functioning sanitation system.[21] Flowers (2018) explains that the soil in Lowndes County is composed mainly of clay, making the installation of wastewater management systems challenging and expensive.[22] As a result, only two towns in the county have centralized water systems, and most residents rely on septic systems to process wastewater. However, the condition of the soil means that septic systems are much more expensive to install than in many areas of the country. Flowers estimates that home septic systems could cost Lowndes residents between $6,000 and $30,000 to install. However, low household incomes and high poverty rates mean that, for many Lowndes County residents, purchasing a functioning septic system for their property is well beyond their financial means. As a result, some homes have systems that are only partially functional, while others lack any septic system at all. Data from the 2014 Alabama Center for Rural Enterprise survey show that the majority of residents deal with septic system problems that at times cause raw sewage to back up into their homes. Residents also frequently encounter raw sewage in their yards and on land throughout the county.

In Lowndes County, census data show that high poverty rates and low median incomes have persisted across generations, demonstrating the ways in which contextual inequity manifests in practice and how intersecting factors can compound harms. In Okeowo (2020), the author profiled several residents of Lowndes County and found that the county's sanitation problems are tied to a large set of issues, including persistent poverty and limited economic issues.[23] For example, lack of adequate sanitation and the widespread presence of raw sewage make any further economic development in the county challenging. Residents are also plagued with health problems connected to poor sanitation, making even the smallest daily tasks challenging, and they face barriers to access health care, which makes it more difficult to seek treatment.

The inequity in Lowndes County is further intensified by other policies: for instance, failure to maintain a functioning septic system is a criminal misdemeanor in Alabama. Winkler and Flowers (2017) find that residents have been arrested and fined for not having

[20] Mark Montgomery and Trevor Logan, "Poor Cybersecurity Makes Water a Weak Link in Critical Infrastructure," Foundation for Defense of Democracies, 2021.

[21] Catherine Coleman Flowers, "America's Dirty Shame: Living amid Raw Sewage," *Anglican Theological Review*, Vol. 100, No. 1, 2018.

[22] Flowers, 2018.

[23] Alexis Okeowo, "The Heavy Toll of the Black Belt's Wastewater Crisis," *New Yorker*, November 23, 2020.

adequate sanitation on their property, adding to the long list of problems that they already face.[24] Arrests become a part of their criminal record and the inability to pay fines creates ongoing legal issues, including potential jail time.

Despite their deep importance to people's lives, utilities that supply water and manage sewage also have weak cybersecurity, as noted in the Cyberspace Solarium report and elsewhere, while they are regularly targeted by malicious actors.[25] There are approximately 52,000 entities that supply drinking water, and 16,000 that handle waste, many of which have limited cybersecurity expertise.[26] It is not just cyber threats that pose risks—recent research identifies Supply Water (along with Provide Public Safety) as the NCFs at the greatest risk of disruption from such climate change events as flooding, drought, and wildfires.[27] However, despite the importance of these utilities to people's well-being and the risks of disruption, they might not appear in the economic lists of SIEs because of their regional role or their limited economic values.

Conclusion

This example from Lowndes County illustrates just one way in which systemic importance can be analyzed from the equity perspective, and more work must be done to refine the analysis and extend it to other examples. This work will include developing a better understanding of the key underserved communities that are overlooked from the economic perspective and need to be prioritized. There are vulnerability and other metrics that help identify these communities, but these indexes are not without their challenges.[28] However, the NCF framework provides a useful structure to identify underserved communities in relation to specific NCFs that will help tie those the underlying considerations that suggest community vulnerability to specific entities.

In addition to identifying the communities that should be prioritized, further work must seek a better understanding of the relationship between the NCFs, the specific entities serving identified communities, and the communities' overall well-being. Much of this research might be regionally specific or geographically contained, though there will likely be linkages

[24] Inga T. Winkler and Catherine Coleman Flowers, "'America's Dirty Secret': The Human Right to Sanitation in Alabama's Black Belt," *Columbia Human Rights Law Review*, Vol. 49, 2017.

[25] U.S. Cyberspace Solarium Commission, 2020.

[26] Montgomery and Logan, 2021.

[27] Michelle E. Miro, Andrew Lauland, Rahim Ali, Edward W. Chan, Richard H. Donohue, Liisa Ecola, Timothy R. Gulden, Liam Regan, Karen M. Sudkamp, Tobias Sytsma, Michael T. Wilson, and Chandler Sachs, *Assessing Risk to the National Critical Functions as a Result of Climate Change*, HSOAC operated by the RAND Corporation, RR-A1645-7, 2022.

[28] Some of the key metrics include the Social Vulnerability Index (Adaptation Clearinghouse, "Social Vulnerability Index (SoVI)," webpage, undated) and the CDC Social Vulnerability Index (Agency for Toxic Substances and Disease Registry, "CDC/ATSDR Social Vulnerability Index," Centers for Disease Control and Prevention, 2022). These have been critiqued in Finucane, May, and Chang, 2021.

across regions. Lastly, the work will require identifying the contextual, distributional, and procedural impediments to achieving equity those populations face with respect to the identified entities.

This work on the equity perspective will help ensure that DHS can meet the Biden administration's commitments to address the needs of underserved populations, and to identify and engage SIEs important for them. This will involve developing relationships and engaging key identities, identifying gaps, and implementing policies and programs to provide needed cybersecurity and other support.

CHAPTER FOUR

Systemic Cyber Risk

Cyberspace has continuously grown in complexity as infrastructure has become increasingly reliant on the cyber ecosystem to function. As cyberspace has become more and more complex, so too have cyber threats. Data breaches have exposed the personal information of hundreds of millions of individuals, causing widespread vulnerability to identity theft.[1] In the years since Solarium, these incidents have only increased. In February 2021, unidentified hackers remotely accessed controls at a water treatment facility in a small Florida town, attempting to increase chemicals in the water to dangerous levels.[2] In May 2021, JBS USA Holdings, Inc., the No. 1 beef producer in the United States and responsible for packing nearly a quarter of domestic beef supply, fell victim to a ransomware attack. The incident resulted in disruptions at all U.S.-based facilities and cost the company $11 million in ransom payments.[3] The attack came just weeks after another ransomware attack led to the shutdown of the Colonial Pipeline, which transports 2.5 million barrels of fuel a day across the East Coast.[4]

In this chapter we discuss data-driven approaches for identifying and prioritizing cyber risk. For the purpose of this discussion, *cyber risk* refers to potential harm resulting from a computer-enabled failure or malicious incident, as determined by the likelihood and associated consequences of such an event.[5] In particular, we consider cyber risk at the *firm* level and the *software* level. That is, we consider variation in cyber risk across different firms (as a

[1] Two such examples are the September 2017 data breach of Equifax which exposed the private information of 147 million individuals and the June 2015 U.S. Office of Personnel Management data theft of background investigation records of 21.5 million individuals. For more on these incidents, see Federal Trade Commission, "Equifax Data Breach Settlement," webpage, February 2022. See also Office of Personnel Management, "Cybersecurity Resource Center: Cybersecurity Incidents," webpage, undated.

[2] Frances Robles and Nicole Perlroth, "'Dangerous Stuff': Hackers Tried to Poison Water Supply of Florida Town," *New York Times*, February 8, 2021.

[3] Jacob Bunge, "JBS Paid $11 Million to Resolve Ransomware Attack," *Wall Street Journal*, June 9, 2021.

[4] William Turton and Kartikay Mehrotra, "Hackers Breached Colonial Pipeline Using Compromised Password," *Bloomberg*, June 4, 2021.

[5] This definition is derived from the simple definition of risk in the DHS Lexicon: "potential for unwanted outcome resulting from an incident, event, or occurrence, as determined by its likelihood and the associated consequences" (DHS, *DHS Risk Lexicon: 2010 Edition*, September 2010).

function of their network configurations, and supply chain), and cyber risk across software applications (as a function of their prevalence, vulnerabilities, and library dependencies).[6]

The approaches developed here align with the NRMC's strategic intent in several ways.[7] For instance, they

- provide the NRMC with "timely access to available data on the risk posture of key information systems" and help "prioritize the risks" (Cyber Defense, sections 1.1.1 and 1.1.2)
- provide awareness of "imminent hybrid, supply chain, and emerging threats and their potential impacts on society" (Critical Infrastructure Resilience and Capacity Building, section 3.1.1)
- ensure strategic risk posture awareness by ensuring that the NRMC "knows the risk postures of agencies with an accuracy and fidelity commensurate with risk to the critical functions of the federal enterprise" (Federal Cybersecurity Governance and Capacity Building, section 3.2.1)
- assist the NRMC to "anticipate, understand, and respond to long-term risks" (Long-Term Risk Management, section 4.2.1).

Cyber Risk of Systemically Important Entities

In the course of this effort, we identified four approaches for understanding cyber risk to SIEs, and we consider how each approach might be suitable for answering different kinds of policy or research questions (i.e., use cases).[8]

First, one approach would be to develop a single aggregated measure of the overall cyber risk of *any given firm* (SIE). A second approach would be to map out a firm's internet-accessible software vulnerabilities to define or assess a firm-level cyberattack surface.[9] Both of these approaches can be applied to any list of SIEs: companies (and their associated supply chain) within a specific NCF, companies within one or more subcomponents of an NCF, or all companies across all NCFs. A third approach would be to focus specifically on the cyber risk of software and IT companies, given their importance in affecting cyber risk for their customers. Finally, a fourth approach would be to assess the risk posed to a company because of its software and IT supply chain. We describe these approaches in more detail below.

[6] Further development might also include computing hardware.

[7] A partial list of NRMC strategic intent goals is listed in Appendix B.

[8] The approaches were developed based on insights developed and collected from previous engagements with this sponsor. There might well be additional approaches to consider, but we begin with the ones listed here.

[9] We consider the attack surface as an organization's internet facing computing systems, each of which might provide an opportunity for a malicious actor to compromise the organization and its data.

1. Enterprise Cyber Risk Metric

A potentially useful capability for a decisionmaker is to create or use a single metric that serves as a proxy for a company's enterprise cyber posture. This kind of cyber risk metric can be constructed by collecting and combining security-related information based on publicly accessible information for a given company. For example, data can be collected by scanning a company's public-facing internet services, collecting information about exposed vulnerabilities, enumerating the configuration (or misconfiguration) of internet services, or counting the volume of spam email emanating from the company's networks. All these data are combined to create a single enterprise score which might reflect, at an aggregate level, a company's overall security posture. More information about the science of network-based cyber risk prediction is available in Liu (2015).[10]

Alternatively, a cyber risk score could be collected (purchased) from a commercial provider, such as Bitsight, Security Scorecard, RiskRecon, or ISS Cyber Risk.[11] These firms sell scores for individual firms and are aggregated across industries as a way for companies to track and compare their cyber risk with that of competitors. These security providers also sell the scores to insurance companies to help them assess and price a company's cyber risk.

Whether custom built, or acquired from third-party providers, each option has advantages and disadvantages. Creating a custom cyber risk metric has the advantage that all data and model parameters are entirely transparent, which affords the ability to incorporate or remove information as necessary. Moreover, weights or transformations can be applied to variables to optimize one approach over another. On the other hand, computing a comprehensive enterprise risk score can be an extremely complicated effort, requiring constant attention and improvement, and so integrating a commercial solution might be preferred. These alternatives are shown in Table 4.1.

TABLE 4.1
Advantages and Disadvantages of Custom Versus Commercial Cyber Risk Scores

Solution Type	Advantage	Disadvantage
Custom-built solution	Data sources and algorithms are transparent and flexible to accommodate new data, and alternative weighting schemes	It is difficult and complicated to acquire all the necessary data and to merge these data into a single risk score in a transparent and objective way
Commercial solution	Commercial providers have already invested a great deal of time and effort into data collection and refinement of the data	Methods and data sources are proprietary and therefore might not be suitable for some decisionmaking practices

[10] Y. Liu, A. Sarabi, J. Zhang, P. Naghizadeh, M, Karir, M. Bailey, and M. Liu, "Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents," USENIX Security, Washington, D.C., August 2015.

[11] See BitSight, homepage, undated; Security Scorecard, homepage, undated; RiskRecon, homepage, undated; Institutional Shareholder Services, homepage, undated.

An important consideration regarding the inferences that are possible with this kind of cyber risk metric (or, indeed, any such metric) involves understanding the relationship between the specific use case (or question a decisionmaker is asking of the data), and the *completeness* of the available data. For example, consider the following two use cases:[12]

1. What are the cyber risk scores for a given list of 50 companies?
2. What are the top 100 companies with the highest cyber risk scores?

The first use case could be described as an exercise in identification, while the second use case is an exercise in prioritization, and they might each be appropriate, depending on the interest. Although this might be a subtle—even obvious—distinction, there is an important implication for underlying data requirements. The first use case can be answered ad hoc—even in real time if the data readily exists in government systems—by polling a third-party information source (e.g., Shodan) with a list of companies. However, the second question requires that the entire data set be available to scan and sort according to the specific request. For a data source like Shodan, the former could be done ad hoc using a free and lightweight application programming interface (API), while the latter requires having access to an entire data set, a proposition which might come at a substantial financial cost, and might require significant human resources to manage it.

2. Mapping a Systemically Important Entity's Cyberattack Surface

Another approach to understanding SIE cyber risk is to consider the exposure of any given SIE to internet-based malicious incidents, i.e., malicious cyber incidents targeting an SIE's publicly accessible internet applications or infrastructure, what we refer to here as a *cyberattack surface*. The greater the number of publicly accessible devices that exist within a given SIE's control, the more vulnerabilities become potentially exposed, and the larger the SIE's attack surface.

This approach begins with a given list of SIEs. For each company, internet scan data can be collected from open-source or commercial companies (e.g., Shodan, Bitsight) to retrieve the list of software vulnerabilities exposed by the company's publicly accessible internet systems.[13] This list of vulnerabilities would include the unique vulnerability identifier (i.e., CVE ID) and a count of the number of instances found. For each unique vulnerability, additional

[12] By *use case*, we refer to questions that a policy maker or a researcher might be interested in, capabilities or questions that a stakeholder might pose of the data, or the information system described in this report.

[13] As described here, a software vulnerability is a weakness or flaw in a software application that could allow a malicious actor to exploit or compromise the information system, for the purpose of destroying, stealing, or otherwise compromising private or commercial information or computing devices. Software vulnerabilities are catalogued by the Mitre Corporation, a nonprofit that developed the Common Vulnerability and Exposures (CVE) method for uniquely labeling and identifying software vulnerabilities (CVE, "Overview," webpage, undated).

data are collected, such as a measure of the vulnerability's severity (i.e., the impact on the information system if the vulnerability were to be exploited), and the likelihood that the vulnerability would be exploited by a malicious actor.

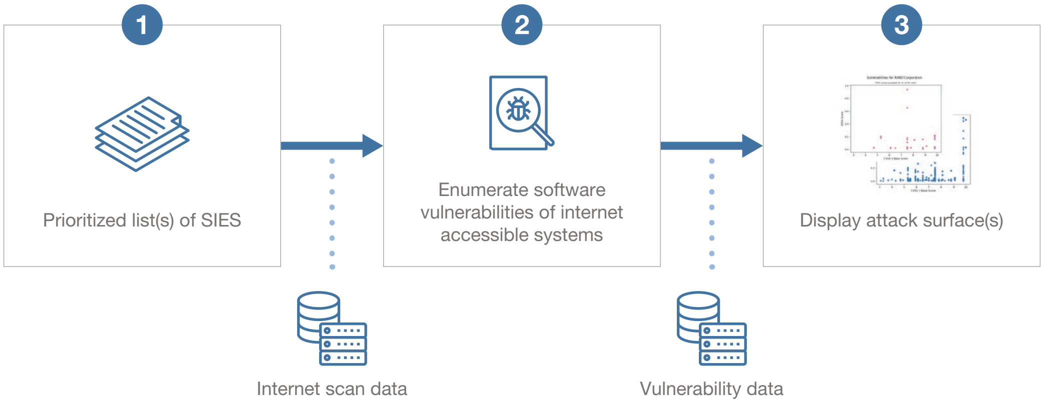
The industry standard method to measure a software vulnerability's severity is the Common Vulnerability Scoring System (CVSS), which is an open standard managed by Forum of Incident Response and Security Teams (FIRST).[14] Similarly, the best source of exploitability information comes from the Exploit Prediction Scoring System (EPSS), another standard which is also part of FIRST. Specifically, EPSS estimates the probability that a vulnerability will be exploited within 30 days from the day the score was generated.[15] CVSS data are collected from the National Institute of Standards and Technology's (NIST's) National Vulnerability Database and EPSS data are collected from the EPSS website.[16]

This process of SIE selection, collecting vulnerability scan data, and visualization of the company's cyberattack surface is shown in Figure 4.1.

A cyberattack surface is one component of an organization's cyber risk profile, but this sort of visualization can be helpful to interpret and compare vulnerability information among and across SIEs. In addition, this information can be used to support the following use cases:

1. compute a cyber threat metric based on the probability of at least one internet-accessible vulnerability being exploited

FIGURE 4.1
Cyberattack Surface


1
Prioritized list(s) of SIES
2
Enumerate software vulnerabilities of internet accessible systems
3
Display attack surface(s)
Internet scan data
Vulnerability data

NOTE: A prioritized list of SIEs (Step 1) is combined with internet scan data to enumerate all the publicly accessible vulnerabilities exposed by those companies (Step 2). Additional data regarding the severity and exploitability of each vulnerability are collected and used to visualize each SIE's cyberattack surface (Step 3).

[14] See FIRST, "Common Vulnerability Scoring System SIG," webpage, undated a.

[15] See FIRST, "Exploit Prediction Scoring System (EPSS)," webpage, undated b.

[16] FIRST, undated b.

2. determine which companies suffer the greatest risk from internet-accessible vulnerabilities
3. determine which companies are vulnerable to a newly discovered software vulnerability.

3. Identify Critical Software/IT Businesses

A third approach for examining SIE cyber risk begins by filtering a business relationship database (e.g., FactSet or Bloomberg) based on an industry code (e.g., FactSet RBICS code) to identify only software and IT companies. Once a filtered list is generated, the companies can then be prioritized based on one or more importance metrics (e.g., centrality, cyber risk). These steps are illustrated in Figure 4.2.

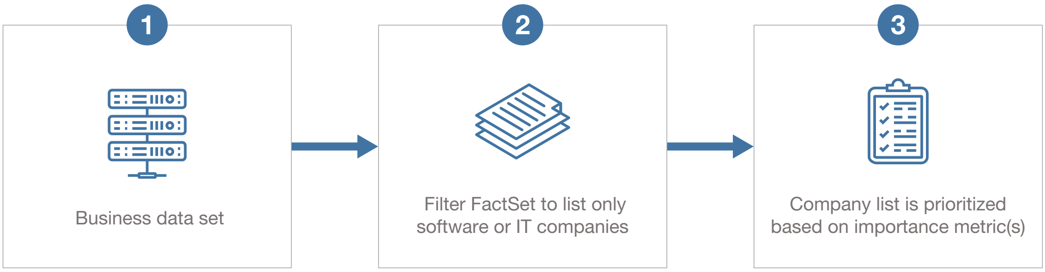
This approach can be used to address the following use cases:

1. A new vulnerability or compromise has been detected in a specific application; which companies are using that software?
2. Which software and IT companies have the greatest cyber risk score?

4. Examining the Software and IT Supply Chain for a List of Systemically Important Entities

The final approach begins with an existing list of SIEs and, using a business data set of supply chain relationships (e.g., FactSet, Bloomberg), identifies those companies' software and IT supply chain connections. This can be done either ad hoc for an individual company, or, if performed for a larger group of companies, results can be aggregated. The purpose of this approach is to identify any specific software or IT suppliers for a given set of critically important companies and consider whether those suppliers represent significant risk to the SIEs. These steps are shown in Figure 4.3.

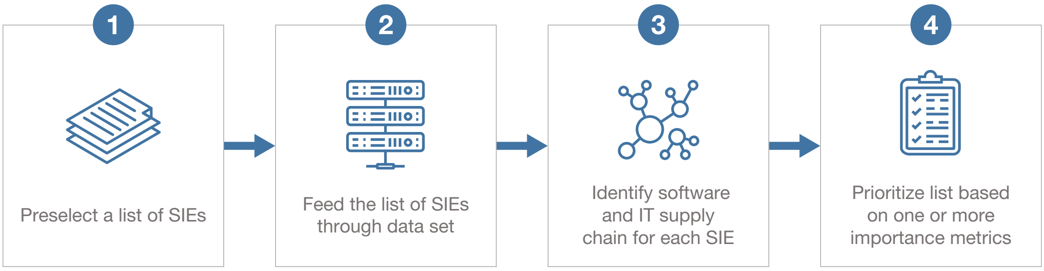
FIGURE 4.2
Identify Software and IT Companies


1
Business data set
2
Filter FactSet to list only software or IT companies
3
Company list is prioritized based on importance metric(s)

NOTE: A business data set (Step 1, such as FactSet) is used to filter companies based on software or IT company industry code (Step 2), which is then prioritized according to one or more importance metrics (Step 3).

FIGURE 4.3

Software and IT Supply Chain for List of Systemically Important Entities


1
Preselect a list of SIEs
2
Feed the list of SIEs through data set
3
Identify software and IT supply chain for each SIE
4
Prioritize list based on one or more importance metrics

NOTE: A list of SIEs (Step 1) is fed through a business data set (Step 2) to identify the software and IT company supply chains (Step 3), which can then by prioritized according to any relevant importance metric (Step 4).

This approach can be used to address the following use cases:

1. For company X, which software and IT companies are part of its supply chain?
2. Which software or IT companies are most important to a given list of SIEs?
3. Within the supply chain associated with a particular NCF, what software and IT entities are the sources of greatest vulnerability?

Summary of Approaches and Use Cases

The collection of approaches and use cases to which they respond are summarized in Table 4.2. Next, we examine cyber risk as examined through the lens of the software supply chain.

Cyber Risks of the Software Supply Chain

Another way of assessing cyber risk is by measuring the overall prevalence of software applications, and exploring how software components are linked (e.g., used or referenced) across disparate applications. For the purpose of this discussion, the following terms are used interchangeably: software application, package, library, software component.

Software Prevalence

A first measure for estimating software risks involves understanding the prevalence of software packages, that is, the number of times a given application is used or referenced. This information can be collected from multiple data sources, and each source might provide different information, therefore supporting different inferences. For example, Veracode/Cyentia and the Core Infrastructure Initiative (CII) are two organizations that have studied the prevalence of software and have published reports describing the most prevalent software

TABLE 4.2

Systemically Important Entity Cyber Risk Approaches

SIE-Based Cyber Risk Approach	Possible Use Cases	Potential Data Sources
Enterprise Cyber Risk Metric	1. What are the cyber risk scores for a given list of 50 companies? 2. What are the top 100 companies with the highest cyber risk scores?	Shodan, Bitsight, Security Scorecard, RiskRecon, ISS Cyber Risk
Mapping an SIE's Cyberattack Surface	3. Compute a cyber threat metric based on the probability of at least one internet-accessible vulnerabilities being exploited 4. Which companies suffer the greatest risk from internet-accessible vulnerabilities? 5. Which companies are vulnerable to a newly discovered software vulnerability?	Shodan, EPSS, CVSS
Identify Critical Software and IT Businesses	6. A new vulnerability or compromise has been detected in a specific application, which companies are using that software? 7. Which software and IT companies have the greatest cyber risk score?	FactSet, Bloomberg Government (BGov)
Examining the Software and IT Supply Chain for a List of SIEs	8. For company X, which software and IT companies are part of its supply chain? 9. Which software or IT companies are most important to a given list of SIEs? 10. Within the supply chain associated with a particular NCF, what software and IT entities are the sources of the greatest vulnerability?	FactSet, BGov

components based on their analysis.[17] These efforts, if representative of the entire software ecosystem, would help inform questions about overall prevalence about a specific software component. This approach can be used to address the following use case:[18]

1. Which software components or libraries are most commonly used?

Software Dependence

The next two approaches for understanding software risk relate to collecting information about how one software component is linked or referenced by another software component. The difference between the approaches—which is an important one—refers to the direction of association.

[17] Frank Nagle, Jessica Wilkerson, James Dana, and Jennifer L. Hoffman, *Vulnerabilities in the Core: Preliminary Report and Census II of Open Source Software*, Linux Foundation Core Infrastructure Initiative, Linux Foundation, February 2020; Cyentia Cybersecurity Research Library, "2020 State of the Software Supply Chain," webpage, September 1, 2020.

[18] Note that although these use cases relate to a software-level analysis, some use cases might be similar to the SIE use cases described above.

We first consider software *dependents*, which refers to all the software packages that *depend on* a given software application, as depicted in Figure 4.4.

This approach provides an intuitive measure of importance of software: the greater the number of components that depend on a package, the more important that package might be, all else being equal. This approach is related to the previous approach but differs in scope. Although the first approach sought to make global inferences about software dependence across all applications, this approach is more focused in that it relates to a specific application or library and can be used to address the following use cases:

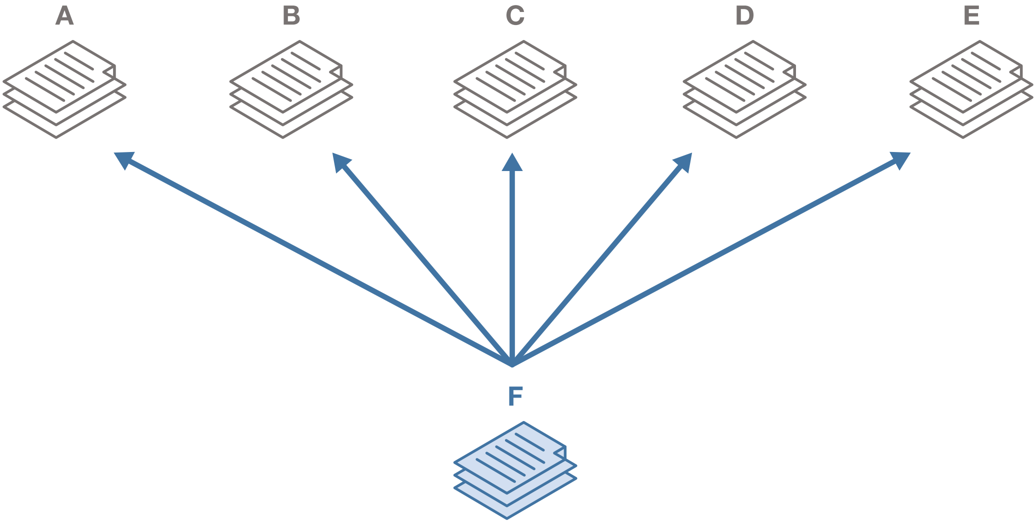
1. A new vulnerability has been identified in package X; how many other software applications depend on this package?
2. Which software packages have the greatest number of dependents?

Software Dependency

Next we consider software dependencies, or what is often referred to as a software bill of materials (SBOM).[19] Although the previous effort sought to identify software packages that depend on a given library, this effort seeks to identify the software packages that are required in order for an application to function properly, as shown in Figure 4.5.

FIGURE 4.4

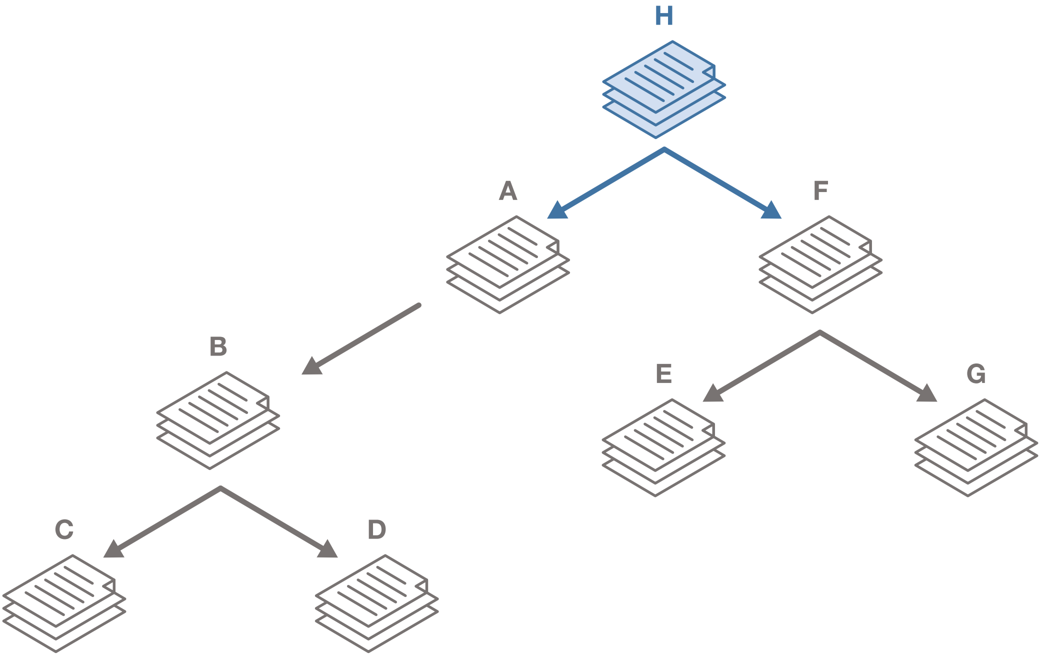
Illustration of Software Dependents Tree



NOTE: This figure depicts how software packages A–E depend on software package F.

[19] See CISA, "Software Bill of Materials," webpage, undated d.

FIGURE 4.5
Illustration of a Software Bill of Materials Dependency Tree



NOTE: This figure depicts how software package H depend on packages A–G.

Mapping out an application's SBOM is a potentially useful way to estimate cybersecurity risk. For example, for each of the dependent packages, the list of known vulnerabilities can be retrieved, along with the vulnerabilities' severity and exploitability. In effect, this collection of vulnerability information would produce the equivalent of a cyberattack surface (similar to what was previously described).

This approach can be used to address the following use cases:

1. How many and which packages does package X depend on?
2. A new vulnerability has been identified in package X; is package Y affected?
3. How many vulnerabilities are potentially embedded in package X?

Summary of Approaches and Use Cases

The collection of approaches and use cases to which they respond are summarized in Table 4.3.

Connecting Business Supply Chain with Software Supply Chain Risks

Consider the following use case: "A critical vulnerability has been newly discovered in a particular software package." What is the national consequence of this vulnerability? And how would we go about answering that question? We outline one approach in the discussion below by using the SIE-based and software-based risk approaches described above. For example, these approaches could be used to estimate the potential impact from the SolarWinds compromise, or the Log4J vulnerability.[20]

First, we present the overall architecture and information flow of this system, as shown in Figure 4.6.

Step 1. The first step involves identifying all companies using the vulnerable software. This could be done in several ways. First, application scanning services like Shodan could be used to identify certain software and components running within a company's public-facing internet space. Alternatively, to identify software purchased by government entities, commercial data sources like Bloomberg's BGov could be used. Information about customers of commercial software providers might also be available from other sources, such as the vendor. Although other data sources could include, for example, GitHub, and Veracode, not all software will be visible through scans of public-facing systems.[21]

Step 2. Once the companies using the vulnerable software have been identified, the list can be prioritized according to one or more importance metrics (e.g., network centrality, economic centrality, cyber risk, equity, NCF participation). This provides an ordered list of the most important companies potentially affected by this newly discovered vulnerability.

Step 3. An optional step is to map the supply chain of suppliers and customers for each company using the software to understand the downstream effects from a disruption or

TABLE 4.3

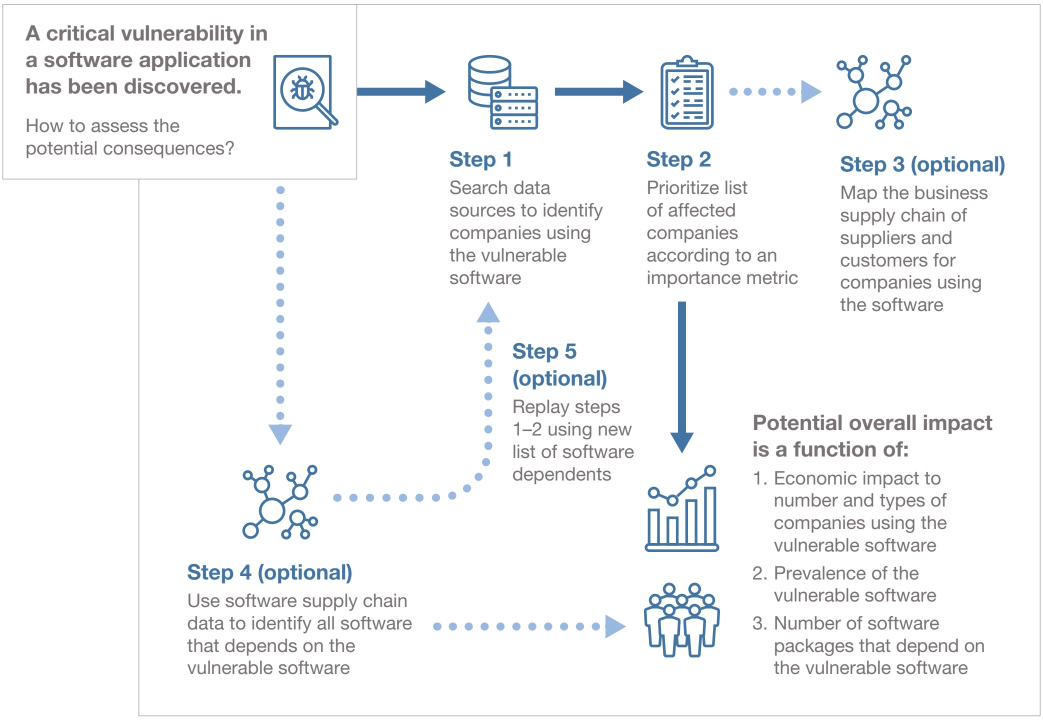
Software Cyber Risk Approaches

Software-Based Cyber Risk Approach	Possible Use Cases	Possible Data Sources
Software Prevalence	1. Which software components or libraries are most commonly used?	Veracode (2020), CII (2020)
Software Dependence	2. A new vulnerability has been identified in package X; how many other software applications depend on this package? 3. Which software packages have the greatest number of dependents?	GitHub, library.io, snyk.io, Synopsys, etc.
Software Dependency (SBOM)	4. How many, and which packages, does package X depend on? 5. A new vulnerability has been identified in package X; is package Y affected? 6. How many vulnerabilities are potentially embedded in package X?	GitHub, library.io, snyk.io, Synopsys, etc.

[20] See Dina Temple-Raston, "A 'Worst Nightmare' Cyberattack: The Untold Story Of The SolarWinds Hack," *NPR*, April 16, 2021; Santiago Torres-Arias, "What is Log4j? A Cybersecurity Expert Explains the Latest Internet Vulnerability, How Bad It Is and What's At Stake," *The Conversation*, December 21, 2021.

[21] Note that external scans of company networks will not reveal information about software or hardware used within the company's network (i.e., behind firewalls or other perimeter devices), and so it is important to recognize the limitations of any data collection.

FIGURE 4.6
Mapping Software Supply Chain


A critical vulnerability in a software application has been discovered.
How to assess the potential consequences?
Step 1
Search data sources to identify companies using the vulnerable software
Step 2
Prioritize list of affected companies according to an importance metric
Step 3 (optional)
Map the business supply chain of suppliers and customers for companies using the software
Step 5 (optional)
Replay steps 1–2 using new list of software dependents
Step 4 (optional)
Use software supply chain data to identify all software that depends on the vulnerable software
Potential overall impact is a function of:
1. Economic impact to number and types of companies using the vulnerable software
2. Prevalence of the vulnerable software
3. Number of software packages that depend on the vulnerable software

compromise. Potential data sources that could be used include FactSet, BGov, Shodan (ISP, Cloud), GitHub, or Veracode.

Step 4. The second major capability is to determine which software packages or applications are dependent on the original vulnerable application. That is, which other applications require the initial vulnerable application to function? Potential data sources that could be used include GitHub, Veracode, snyk.io, or libraries.io.

Step 5. The fifth (optional) step is to use software supply chain data to identify all companies that are using the vulnerable software. Assume that all software packages that depend on the vulnerable software are themselves vulnerable. Given this larger list of software, return to step 2 to identify all companies using the vulnerable software.

Together, the information collected from step 2 and step 4 provides distinct measures by which to evaluate the relative consequence of a vulnerable software application, such as the following:

- a measure of economic impact for all companies using the vulnerable software (or its dependents)
- a measure of prevalence of the vulnerable software, potentially across one or more NCFs
- the number of software applications that depend on the vulnerable application.

Limitations in Current Approaches to Understanding Systemic Cyber Risk That Might Be Addressed Through Future Work

This research uncovered several limitations, each of which could be addressed in future research.

Linking Company Names Between Disparate Data Sets

In the course of this project, we needed to link data between multiple data sets, indexed by company name. Although there exist some industry conventions for uniquely identifying companies (e.g., gvkey in COMPUSTAT), this approach is not universally adopted across all data sources. Indeed, in one test case of a sample of more than 9,000 company names, we were only able to uniquely link about 50 percent of them, even after performing minor data cleaning such as removing trailing words (e.g., ", LLC" or ", Inc."). Although more improvements could be made, a complete solution (i.e., one that matches 100 percent of company names) would take considerable effort. This is one example of a larger namespace problem that we will likely encounter in future work and anytime one data set needs to be merged or linked with another data set by company name.

Limitations with Mapping Software Supply Chain Dependents and Dependencies

The recent vulnerability identified in a popular open-source package (Log4J) raised concerns about the prevalence, and therefore risk, of open-source software used throughout the United States. In addition, there is a movement across the software industry and the U.S. government to incentivize software vendors to create and disclose the software ingredients in their products, otherwise known as a SBOM.[22] This is important because very little is known about the scale and scope of software dependencies, and there is no clear way to acquire sufficient information about the extent and depth of software dependencies.

Thankfully, the problem is not entirely unsolvable. The information required to map software dependents and dependencies is knowable—it exists among the collection of software repositories and related companies. It might be difficult to find and would require a great deal of work, but it does exist. Furthermore, the problem set is bounded; that is, the list of dependencies and dependents for any software package is finite. For example, it is possible to retrieve some information about software dependencies for some packages. GitHub provides an API that can be used to retrieve dependency information for a given package, one dependency layer at a time. Although helpful, this approach would clearly not scale to produce sufficient data for all software repositories hosted on GitHub. Furthermore, retrieving dependent information is less automated. Although some information can be scraped manually,

[22] See EO 14028, "Executive Order on Improving the Nation's Cybersecurity," May 12, 2021.

there is currently no API for this feature. And although GitHub might be a leading repository for open-source software, it is just one among many software ecosystems.

Moreover, although one package might reference another package, it might not actually execute any of that package's functions. Therefore, the existence of a dependency does not automatically infer a vulnerability (and therefore increase risk). This reduces the scope of the problem, but also increases the complexity of accurately identifying it. In addition, it could be the case that dependencies lose effect after enough degrees of separation. However, the true effect is worthy of further research.

Two private-sector efforts have recently emerged to help improve the open-source software ecosystem. First, Scorecard v4 was announced in late 2021.[23] This is a partnership between Google and GitHub to provide a GitHub tool than scans a repository and identifies configurations or workflows that might pose a risk to the integrity of the code, e.g., untrusted code insertions. Next, Alpha-Omega is a software project that was announced in February 2022, as a partnership between Google and Microsoft.[24] It purports to be a combination of code audits, threat modeling, automated software vulnerability scanning, and remediation support for select open-source software packages. For example, it might determine that a particular repository allows unauthenticated code modifications, or that it exposes a default username and password.

Augmenting Business Supply Chain Data

Although FactSet identifies many different kinds of business relationships between companies, there might still be the opportunity to augment this information with additional business relationship data. For example, in addition to software and vulnerability data by firm, Shodan collects information about which ISPs and cloud services are used by each company in its database. Currently, these data are not being integrated into FactSet, but future work could enable this.

Connecting Web Component Data with Open-Source Packages

Currently we do not connect (link) web component data from Shodan with open-source package data from GitHub. However, doing this would certainly provide new insights, and provide another way to identify new risks.

[23] See Jonathan Greig, "Google Announces Scorecard V4 in Partnership with GitHub and OpenSSF," *ZDNet*, January 19, 2022.

[24] See Paul Sawers, "Google and Microsoft Back the Alpha-Omega Project to Bolster Software Supply Chain," *VentureBeat*, February 1, 2022.

Augmenting Software Data

Additional metadata could also be collected to reveal further insights, including the following:

- the **number of contributors** provides a measure of the complexity of the project, and perhaps the resilience of the repository to changes in developers
- the **organizational structure of the project** is also a measure of the complexity of the project and of the resilience of the project to organizational changes
- **foreign ownerships** might be useful when tracking organizational leaders or foreign influence
- the **number of, or most recent, commits** provides a sense of how frequently updated and current the package is, for example, whether it is at end-of-life or actively updated and patched
- the **number of branches (forks) in the project** could provide an idea of how the repository is used, maintained, or depended on by others outside its direct contributing community.

Resilience and Recommended Security Controls

The discussions provided here address opportunities to identify cyber risk across business and software supply chains. However, they do not address or reflect any measure of resilience (the degree to which a firm could withstand and recover from a cyber incident), or the appropriate security controls one or more organizations could apply to reduce or minimize the risk.

CHAPTER FIVE

Future Research

Significant work remains in developing concepts and modeling approaches for systemic risk to critical infrastructure, advancing the NRMC's incorporation and stewardship of data sets for analysis and visualization, maturing the SIE Program Office processes and procedures for analysis and outreach, and advancing HSOAC's SIAM to reflect emerging perspectives for prioritization—including public health and safety, national security, equity, and others. This chapter addresses research questions and analysis needs on the horizon—work that would help advance the NRMC's risk reduction mission by identifying and addressing risks before they are leveraged to disrupt the nation's critical infrastructure.

Although the SIE concept has been motivated by the challenges of systemic cyber risks, its potential extends across all sources of systemic risk. This chapter details several potential areas for research and analysis against four topics: SIE concepts and modeling approaches; data management methods for expected increase analytic input data; program governance and processes to ensure SIE as a sustainable program; and continued refinement (and transfer) of the prototype SIE analytic platform (SIAM).

Advance Systemically Important Entity Concepts and Modeling

Research question: How should the identification and prioritization of SIEs be extended beyond today's understanding to capture nonobvious dependencies that could comprise a list of "other systemically important entities" (OSIEs) that, although potentially smaller than SIEs, have significant impacts for their size?

Perhaps the most extensive area for future research is the need to advance SIE concepts and modeling approaches. There is broad understanding that the existing focus on economic modeling and interfirm business relationships is innovative, but only a first step toward understanding the complex issues associated with SICI. Additional modeling methods, such as those suggested by the Secure Tomorrow toolkit, could be extended to include scenario-based planning, tabletop games to aid in stakeholder decisionmaking before crises, and the use of agent-based simulation to probe emerging phenomena and potential causes and effects from disruption to SIEs.

Additional perspectives—beyond economic interdependency—could include understanding the systemic risk posed by critical infrastructure disruption to the defense industrial base, public health and safety, national security, and equity perspectives for SIE prioritization and resource planning. Additional perspectives are likely, and SIE concepts are likely to evolve in the future in response to external events. In the short term, there might be a need to understand the systemic risks to SIE subsidiaries and assets, where relationships among assets could occur at levels not revealed through existing interfirm network analysis. In addition, this research could surface region-specific vulnerabilities to critical infrastructure far beyond well-known issues such as hurricanes and oil facilities.

There might be a rich vein of research to be mined regarding the identification of generic indicators that can characterize a firm as an SIE. Today's approach models the present; the development of predictive indicators might help understand the future. The portfolio of NRMC initiatives, from 5G to quantum computing, illustrate just a part of the changing face of technology and business—SIE indicators might help the NRMC recognize emerging importance to allow for early identification and engagement. These indicators and other research might help the NRMC explore analytic concepts that contribute to understanding emergent risk and over-the-horizon threats, allowing for adaptive response and engagement.

From a cyber perspective, there is a need to understand the dependencies that exist not just among businesses and common software but also within common firmware, hardware, the web infrastructures that are shared, and even physical infrastructure that affects cyber assets (e.g., data centers and undersea cables). This research task could include collaboration with other researchers—some of whom are analyzing the structural risks to the global internet—to develop methods for identifying domestic and international internet bottlenecks and vulnerabilities.

Develop Data Management Methods and Plan for Analytic Input Data

Research question: How should additional data sets be incorporated into the NRMC modeling environments to enable long-term and short-term understanding of emerging systemic risk (e.g., cyber, climate, pandemic)?

The innovative use of both public and proprietary data regarding SIEs aided considerably in the analytic results for this study. Future research might build on this success by exploring ways to include and manage multiple data sets for long-term use supporting analyses. This could include CISA establishing a relationship with multiple industry-specific data sets from various vendors and developing methods to establish a level of data interoperability among them that allows for multi-source analyses that support both annual SIE review and real-time tasking such as those driven by events in 2021–2022. One challenge will be the need to reconcile company namespace across multiple data sets—critical to link or integrate across potential analytic input data. This effort could also consider the incorporation of nontraditional data sets for including and insight regarding emerging threats (e.g., historical threat data for insight into the development of risk indicators). One vision could be a stable data set portfolio for CISA SIE analysis across the agency.

Advance SIE Program as a Sustainable Program

Research question: What should be done to ensure the SIE program is successful over time?

The SIE Program is envisioned as an annual collaborative effort across Sector Risk Management Agencies and other stakeholders. The probability of success for the program might lie in establishing such mechanisms as repeatable methods and processes; advising any need for additional guidance or authorities to enable emerging policy; and developing the operational support scaffolding (such as training, surge support training, and other primer or familiarization materials). Following the events of 2020–2022, it is likely the SIE program will take on additional quick-turn tasking in the future. Developing a repeatable approach to SIE analysis and communication might be critical to program success in this context. Research would capture lessons from the coronavirus disease 2019 (COVID-19) response, data calls following high-profile cybersecurity events, and interagency engagement during early 2022. Finally, the program will likely be required to develop outcome and output metrics for its performance, with a key outcome metric framed as overall risk reduction for SIEs. The metric itself would rely on the development of risk reduction metrics for these SIEs, so progress can be measured and reported for future resource planning.

Refine the Systemic Importance Analytic Model

Research question: How should the SIAM be refined, extended, or replaced so that it becomes a valued tool for SIE visualization and communications?

The SIE analytic platform (SIAM) will be central to future research as a visualization and communications tool for use by the National Infrastructure Simulation and Analysis Center and Analytic Division stakeholders. Therefore, the long-term vision is to transfer the tool into the NRMC Modeling Capability Transition Environment with appropriate user guidance, training materials, and initial operating capability and fully operational capability glidepaths. The status as prototype captures the tool as a proof of concept—as ever, caution should be employed in moving a prototype to production. Instead, a technical review informed by documented requirements should be considered to determine how much of the prototype code will be useful for a production tool. This effort involves documenting the emerging requirements and use cases for SIE analysis and visualization, updates to the reflect progress with ongoing NCF decompositions, and incorporation of data interoperability methods discussed above under Data Management.

APPENDIX A

Systemic Importance Analytic Model

Overview or “Welcome” Tab

The current version of the SIAM is meant to help users to identify SIEs whose disruption is not just highly consequential but systemically important to one or more NCFs.

The SIAM allows users to understand the potential risk of an SIE along multiple types of measures, including

- **Network Centrality:** A measure of connectivity to other entities, (i.e., potential to propagate losses upstream and downstream in business networks [calculated at the entity level])
- **Economic Measure:** A measure of an entity’s size denoted by their sales revenue attributed to the specified NCF (calculated at the entity/NCF level)
- **NCF Dominance:** The number of unique NCFs an entity is part of (calculated at the entity level)
- **Revenue Percentage:** The percentage of an entity’s sales revenue attributed to the specified NCF, floored at 0 and capped at 100 in the tool (calculated at the entity/NCF level).

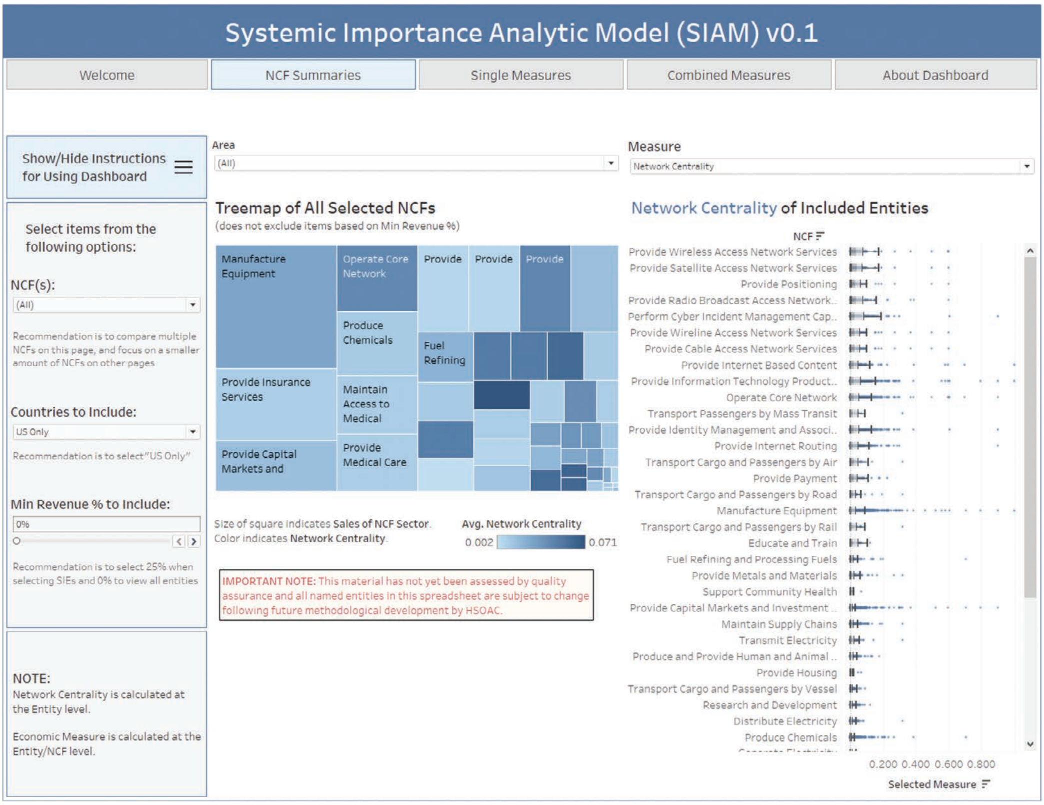
The tool allows users to look at any of the first three measures individually or to view multiple measures together, and it allows individuals to explore the data in the way that works best for them.

“NCF Summaries” Tab

A user can view summaries at the NCF level by either Network Centrality or Economic Measure and can be limited by Area. This might help determine which NCFs to prioritize.

The tree map in Figure A.1 shows the total sales of NCF Sectors, represented by the size of the square, and Network Centrality, represented by color. This chart allows the user to see that the Manufacture Equipment NCF has the largest Economic Measure (as represented by sales), and Provide Wireless Access Network Services has the highest Network Centrality. The chart on the right shows each selected NCF with a dot representing each entity associated with that NCF. This allows the user to see where the bulk of the entities occur for either selected option of Network Centrality or Economic Measure.

FIGURE A.1
Systemic Importance Analytic Model Dashboard



Systemic Importance Analytic Model (SIAM) v0.1

Show/Hide Instructions for Using Dashboard

Select items from the following options:

NCF(s): (All)

Recommendation is to compare multiple NCFs on this page, and focus on a smaller amount of NCFs on other pages

Countries to Include: US Only

Recommendation is to select "US Only"

Min Revenue % to Include: 0%

Recommendation is to select 25% when selecting SIEs and 0% to view all entities

NOTE:
Network Centrality is calculated at the Entity level.

Economic Measure is calculated at the Entity/NCF level.

Area: (All)

Measure: Network Centrality

Treemap of All Selected NCFs
(does not exclude items based on Min Revenue %)

Size of square indicates Sales of NCF Sector.
Color indicates Network Centrality.

IMPORTANT NOTE: This material has not yet been assessed by quality assurance and all named entities in this spreadsheet are subject to change following future methodological development by HSOAC.

Network Centrality of Included Entities

NCF
Provide Wireless Access Network Services
Provide Satellite Access Network Services
Provide Positioning
Provide Radio Broadcast Access Network..
Perform Cyber Incident Management Cap..
Provide Wireline Access Network Services
Provide Cable Access Network Services
Provide Internet Based Content
Provide Information Technology Product..
Operate Core Network
Transport Passengers by Mass Transit
Provide Identity Management and Associ..
Provide Internet Routing
Transport Cargo and Passengers by Air
Provide Payment
Transport Cargo and Passengers by Road
Manufacture Equipment
Transport Cargo and Passengers by Rail
Educate and Train
Fuel Refining and Processing Fuels
Provide Metals and Materials
Support Community Health
Provide Capital Markets and Investment ..
Maintain Supply Chains
Transmit Electricity
Produce and Provide Human and Animal ..
Provide Housing
Transport Cargo and Passengers by Vessel
Research and Development
Distribute Electricity
Produce Chemicals
0.200 0.400 0.600 0.800
Selected Measure

"Single Measures" Tab

A user can view summaries for "Network Centrality," "Economic Measure," or "NCF Dominance," and each sheet will create a list of entities prioritized by the selected measure. Rather than looking at all of the entities or even the top 100 entities or some other prespecified number, there are graphs available (on the Network Centrality and Economic Measure sheets) that show the specified measure for each included entity sorted by rank, and see if there is a noticeable knee-in-the-curve to inform how many entities should be included or what measure limit might make sense to use as a cutoff.

To avoid including a large entity that has only a small impact on a given NCF, we suggest applying a 25-percent NCF Revenue threshold when using only Network Centrality, because that measure is calculated at the entity level.

“Combined Measures” Tab

A user can combine both Network Centrality and Economic Measure to see which entities might be high on one or both, and each sheet will create a list of SIEs. The “Quadrant” sheet allows the user to select two measures (likely Network Centrality and Economic Measure) and to create an SIE List that includes any entities over a specified threshold on either measure. The “Flux Capacitor” sheet allows the user to combine Network Centrality and Economic Measure (a user can select how much to weight each measure) and to create an SIE list prioritized by the new Combined Measure. There is also a knee-in-the-curve graph to help inform entity cutoff counts using the Combined Measure.

“About Dashboard” Tab

Lastly, there are “Methods” and “Glossary” sheets available to provide additional details about the tool and the underlying data.

APPENDIX B

Cybersecurity and Infrastructure Security Agency Strategic Intent and National Risk Management Center Missions and Objectives

The HSOAC study supported the overall CISA mission as detailed in this appendix.

CISA defines its overall mission to "lead the national effort to understand and manage cyber and physical risk to our critical infrastructure" and its vision as a "secure and resilient critical infrastructure for the American people."[1]

CISA's strategic intent ("Defend Today, Secure Tomorrow") is reflected through its statement of CISA Strategic Goals.[2] Goals specific to the NRMC are as follows:

- 2.4. Long-Term Risk Management: Long-term risks are addressed through collaborative risk management across the community.
  - 2.4.1. Analysis Planning and Innovations: CISA anticipates, understands, and responds to long-term risks.
  - 2.4.2. Secure by Design: Systems, assets, and services are designed with the security and resilience of national critical functions in mind.
  - 2.4.3. National Workforce: There is an appropriate supply of security professionals for the national demand.

[1] CISA, *Resource Planning Guidance: FY2023–2027*, December 2020b, p. 1.

[2] See CISA, "CISA Strategic Intent," webpage, undated a.

To achieve these goals, the NRMC is responsible for the priorities specified for CISA's Mission Area 4:[3]

- 4.1. Strategic Risk Analysis: Improve cross-cutting risk analysis capabilities.
- 4.2. Emerging Threat and Strategic Risk Coordination: Improve CISA's capability to address foreign influence and other emerging and strategic risks through policy, planning, and coordination actions.

In practice, the NRMC leverages

> sector and stakeholder expertise to identify the most significant risks to the nation, and to coordinate risk reduction activities to ensure critical infrastructure is secure and resilient both now and into the future. . . . The NRMC creates an environment where government and industry can collaborate and share expertise to enhance critical infrastructure resiliency within and across sectors.[4]

The HSOAC study directly supported the NRMC's National Infrastructure Simulation and Analysis Center, which is "focused on building advanced analytic tools that provide comprehensive, quantitative, and actionable information to enhance CISA's understanding of how to manage risks from a variety of threats to the Nation's critical infrastructure."[5]

[3] CISA, 2020, p. 15.

[4] CISA, "National Risk Management," webpage, undated c.

[5] CISA, "National Infrastructure Simulation and Analysis Center," webpage, undated b.

APPENDIX C

Cyber Data and Software Dependencies

Summary of Data Sources

To inform these approaches, numerous data sources can be used and analyzed (see Table C.1). For example, answering particular research or policy questions requires examining data concerning business relationships. That is, data connecting a given firm with its suppliers (i.e., companies that provide input resources), and customers of that company.

Three-Layer Software Bill of Materials for Apache-Log4J

Figure C.1 shows the software dependency network for Apache-Log4J package. These data were collected using Gitau's API, recursively for three layers of package dependencies. The visualization was done in Gephi.

This SBOM consists of 2,373 software libraries (nodes) and 10,525 references (edges). The color groups represent disparate communities of connected libraries. Overall, we identified one package with almost 500 connections, and 28 packages were linked to more than 100 other software packages.

We also applied a centrality measure to the software packages, which are listed in decreasing importance in Table C.2.

TABLE C.1
Cyber Data Sources

Data Category	Data Source
Enterprise cyber risk score	Bitsight, Cyence, Security Scorecard, etc.
Firm-level cyber data	Shodan, Bitsight, Censys, etc.
Business supply chain data	FactSet, Bloomberg
Software prevalence	Veracode/Cyentia, CII
Software dependence	GitHub, library.io, snyk.io, Synopsys

FIGURE C.1

Dependency Network of Apache-Log4J

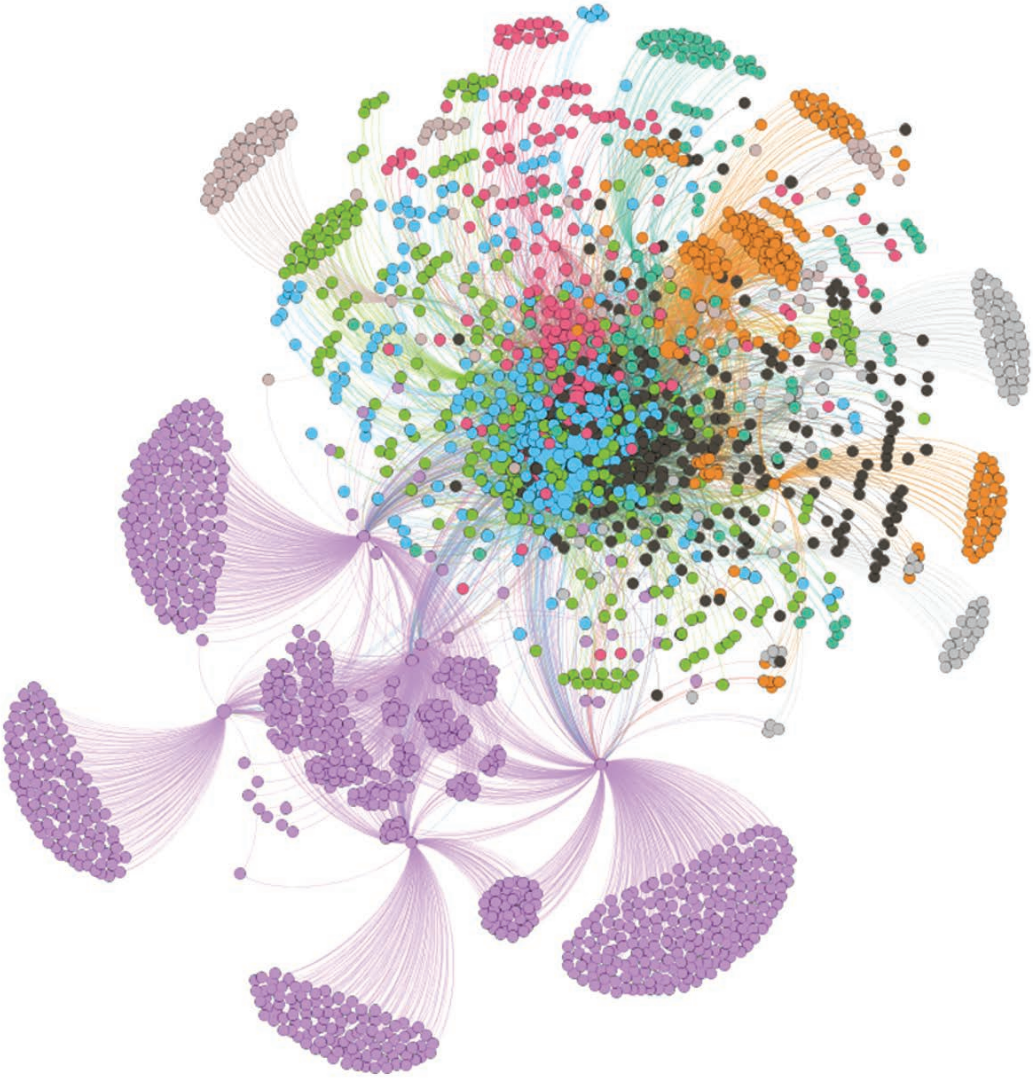


TABLE C.2
Apache-Log4J Dependencies

Importance Rank	Software Publisher/Package	Connections
1	Apache/maven-plugins	399
2	Junit-team/junit4	312
3	Apache/maven-surefire	316
4	Sonatype/sonatype-bundle-plugin	188
5	Apache/maven	183
6	Sonatype/plexus-utils	91
7	Apache/maven-doxia	53
8	Mojohaus/build-helper-maven-plugin	137
9	Mojohaus/animal-sniffer	69
10	Mvonrenteln/hamcrest-java-maven-integration	62

APPENDIX D

Analytical Support for the Cybersecurity and Infrastructure Security Agency

Context

The focus on national critical infrastructure has intensified in recent years, with the growth of criminal economies focused on global ransomware attacks and cyberattacks by various nation-state actors generating interest in and resources to CISA's and NRMC's missions. The associated concerns were arguably exacerbated by the global pandemic of COVID-19, unrest in Europe, and the continuing dangers posed by a warming planet. Therefore, part of the tasking to HSOAC was to aid the NRMC in adapting to an evolving set of mission-related tasks associated with expected EOs, draft legislation, and a new administration's vision for CISA and the NRMC. In addition to accelerating the delivery of certain SIE lists in response to interagency tasking in February and March 2022; we also provided real-time analytic support for the NRMC's response for EO 14028 (May 2021), and ongoing coordination with the National Cyber Director and the White House National Security Council's Senior Director for Resilience and Response.

This chapter briefly summarizes the tasking and insights developed as part of this adaptive response team's work.[1]

Support to the Cybersecurity and Infrastructure Security Agency on Executive Orders on Improving Critical Infrastructure

The 2021 Executive Order on Improving the Nation's Cybersecurity (EO 14028) tasked CISA to "identify and make available a list of categories of software and software products in use

[1] Portions of this chapter were provided earlier as a memorandum for the study sponsor.

or in the acquisition process meeting the definition of critical software."[2] NIST developed an initial definition of "critical software" across 11 specific categories,[3] and CISA requested HSOAC support to develop approaches to identify specific "top" critical software products in accordance with NIST definition. On July 12, 2021, HSOAC submitted a quick-turn memorandum that provided CISA with

1. top products by NIST categories
    a. A proposed approach on how to develop lists of specific software products representing the "top" products for each category
    b. A proof-of-concept initial set of "top" products defined by "market share" and "government wallet share"
2. descriptions of NIST-identified categories of critical software under EO 14028.

This section will briefly lay out the contents of the HSOAC memorandum.

The Top Products, by National Institute of Standards and Technology Category

One approach to identifying top products in any given market is to assess its market share, where market share is defined for any given company or product as equal to its total revenue divided by the sum total revenue of all companies or products in its market segment.

We find two broad reasons why market share might be misleading for identifying the top critical software products. First, the 11 NIST categories do not correspond to "markets" in the traditional sense. For example, software markets defined under Standard Industrial Classification codes divide (e.g., "prepackaged software," "computer programming services," "computers and computer peripheral equipment and software," "business consulting services," and "magnetic and optical recording media") and NAICS (e.g., "software publishers," "software and other prerecorded compact disc, tape, and record reproducing," "custom computer programming services," and "other computer related services") have no relationship to the newly defined NIST categories. Consequently, NIST categories are not markets amenable to external market analysis. Although assessments of each of the 11 markets could be made using product-specific revenue totals in concept, this information is likely proprietary, adding to limitations for data collection on private company sales in practice. Second, if new assessments were to be made of the markets for software under NIST categorization, software use by federal agencies could, in all likelihood, vary significantly from that in the broader market. For both reasons, assessments of market share could be misleading for software use by federal agencies.

[2] EO 14028, 2021.

[3] NIST definition of critical software under EO 14028 is available at NIST, "Definition of Critical Software Under Executive Order (EO) 14028," October 13, 2021.

Therefore, an estimation of wallet shares rather than market share might provide a more-accurate representation of top products for federal agencies, where a given company's/product's federal wallet share is equal to the value of its total sales to the federal government divided by the sum total federal spending on all companies/products in its market segment.

We took the first step toward this assessment by binning software products into the 11 NIST categories. An estimation of wallet share can search across federal contracts—current and upcoming—to assess federal spending on each product across agencies. The total value of each product over the sum of total spending in each respective software category will yield a value of each product's federal spending wallet share. Using this metric, we can develop the top products as the top five in each category based on federal investment.

Caveats

Assessments of market share and wallet share have their limitations. First, the approach provides merely a first approximation of top products by revealing the overall installation base (use) of the software, relative to competing software products. All things equal, the top products are the specific products with the greater wallet share in a given software category.

However, this approach provides no information about the use of that software within the organization, such as the organization's dependence on the software for maintaining business operations, or whether the organization has substitute methods for accomplishing the same tasks if the software experiences a service disruption. For example, although the Google Chrome browser might be the most commonly used web browser,[4] there are many other easily available substitutes in the event of a Chrome service disruption making the browser unavailable. However, substitutes might not alleviate the risk of compromise.

Moreover, although the wallet share approach can be used to rank order the top software products (again, for a given category), additional information would be needed to determine which ones are "most" important. That is, it is conceivable that the top N number of products might be, collectively, considered to be top products because of their aggregate market share. For example, consider products A, B, C, with wallet share of 50 percent, 40 percent and 10 percent, respectively. A and B are top products for the purpose of this exercise, rather than only product A. Furthermore, the approach is not able to provide information about the relative importance of products across product categories. For example, market share is not a good approach for determining whether the most commonly used product in the "Operating System" category is more (or less) important than the most commonly used product in the "Web Browser" category.

The memorandum provided a brief overview of the top five products for several NIST categories, as reflected in market share (notwithstanding key caveats discussed above) (see Table D.1).

[4] StatCounter, "Browser Market Share Worldwide: May 2021," webpage, undated.

TABLE D.1

Top Five Products by Market Share and Alignment to Federal Wallet Share Category (NIST)

Category	Type of Product (NIST)	Top Five Market Leaders (Vendor)	Market Share (%)	Federal Spending to Date with Vendors as Contractor (Versus All Contractors)
Operational Monitoring and Analysis	Security information and event management systems	Splunk (Splunk)	58.74	$8M ($5B)
		IBM QRadar (IBM)	8.37	$2M ($2M)
		LogRhythm (LogRhythm)	6.06	$816K ($44M)
		Rapid7 (Rapid7)	5.92	$50K ($163M)
		ArcSight ESM (Micro Focus Int'l)	3.19	N/A ($3M)
Network Control	DNS resolvers and servers	GoDaddy DNS (GoDaddy.com)	49.87	N/A ($35K)
		Cloudflare DNS (Cloudflare)	15.58	N/A
		Amazon Route 53 (Amazon.com)	4.45	N/A
		Google Cloud DNS (Google)	3.44	N/A ($1M)
		Google Domains (Google)	3.44	N/A
Network Control	Virtual private network (VPN) software	Cisco AnyConnect (Cisco)	28.96	N/A ($128M)
		Cisco VPN (Cisco)	25.80	N/A ($147M)
		Juniper VPN (Juniper Networks)	11.27	N/A ($910K)
		Citrix Gateway (Citrix)	8.67	N/A ($23K)
		OpenVPN (OpenVPN)	3.38	N/A
Network Monitoring and Configuration	Network Management systems	SolarWinds (SolarWinds)	12.48	Insufficient data
		Juniper (Juniper Networks)	12.34	
		Wireshark (Riverbed)	11.99	
		Automate	10.28	
		Novell (NovoPath)	5.52	

NOTE: These results come from text queries for product keywords across government contracts within the GovWin database filtered by the software vendor as the contractor. A different approach would be needed to identify software bundled under other contracts by system integrators which would likely lead to larger numbers. For example, although our search finds $8 million in contracts to Splunk, we also find several billion dollars' worth of contracts that mention Splunk. B = billion; K = thousand; M = million; N/A = not available.

Our approach used the following steps:

1. map subcategories identified by NIST to market analyses. Although NIST categories do not themselves define software markets because they are largely aggregations of several software types, they do provide nearly atomic market definitions under "types of products" in the table describing critical software.
2. review publicly available market analyses that claim to present top products for the software types. We present several of these "top market share" examples in Table 1.2. based on the Datanyze market analysis platform.[5] HSOAC has not independently verified the information presented there and includes the information here as a proof of concept for developing market share as one criterion for determining "top products." As found with other data vendors, the categories used by Datanyze align neatly with only a portion of the software types of interest to this study.
3. use text enabled search through the GovWin government contracts database to identify the total federal spending to date associated with product specific keywords and contracts filtered by vendor name. We identify two values; first, a dollar value where the contractor on record is the vendor of each software product and second, a dollar value for all mentions of the product name across all contractors. The estimates of total federal spending to date shed light on the wallet share approach by estimating the numerator.

Support to the Cybersecurity and Infrastructure Security Agency on Priority Systemically Important Entities

On February 22, 2022, the NRMC requested a list of priority SIEs associated with Phase 1 and Phase 2 NCFs outlined in Table D.2 and any information HSOAC could provide on the remaining NCF phases. Additionally, the NRMC requested insight on entities tied to NCFs associated with the following sectors: Information Technology, Communication, Energy (including nuclear), Finance, Transportation (including pipelines), Public Health, and the Defense Industrial Base. On February 23, 2022, HSOAC submitted a quick-turn memorandum that provided CISA with initial SIAM results, the NCFs they include, supporting methods, and limitations.

Initial Systemic Importance Assessment Model Results

The systemic importance of an entity is a condition of its size, interconnectedness, and substitutability that increases the potential impact of its disruption or failure on others within the system or beyond, leading to the potential of significant cascading impacts. Although

[5] Datanyze, "About," webpage, undated.

TABLE D.2

National Critical Function Analytic Phases

Phase	NCF
Phase 1	Distribute Electricity
	Generate Electricity
	Provide Medical Care
	Supply Water
	Transmit Electricity
Phase 2	Manage Hazardous Materials
	Manufacture Equipment
	Operate Core Network
	Produce Chemicals
	Provide Cable Access Network Services
	Provide Identity Management and Associated Trust Support Services
	Provide Information Technology Products and Services
	Provide Internet Based Content, Information, and Communication Services
	Provide Internet Routing, Access, and Connection Services
	Provide Metals and Materials
	Provide Positioning, Navigation, and Timing Services
	Provide Radio Broadcast Access Network Services
	Provide Satellite Access Network Services
	Provide Wireless Access Network Services
	Provide Wireline Access Network Services

Table D.2—Continued

Phase	NCF
Phases 3–4 (NCFs included in SIAM 2.0)	Educate and Train
	Exploration and Extraction of Fuels
	Fuel Refining and Processing Fuels
	Maintain Access to Medical Records
	Maintain Supply Chains
	Manage Wastewater
	Perform Cyber Incident Management Capabilities
	Produce and Provide Agricultural Products and Services
	Produce and Provide Human and Animal Food Products and Services
	Provide Capital Markets and Investment Activities
	Provide Consumer and Commercial Banking Services
	Provide Funding and Liquidity Services
	Provide Housing
	Provide Insurance Services
	Provide Payment, Clearing, and Settlement Services
	Provide Wholesale Funding
	Research and Development
	Store Fuel and Maintain Reserves
	Support Community Health
	Transport Cargo and Passengers by Air
	Transport Cargo and Passengers by Rail
NCFs not included	Conduct Elections
	Develop and Maintain Public Works and Services
	Enforce Law
	Operate Government
	Prepare for and Manage Emergencies
	Preserve Constitutional Rights
	Protect Sensitive Information
	Provide and Maintain Infrastructure
	Provide Public Safety
	Provide Materiel and Operational Support to Defense

HSOAC's research into SIEs was ongoing, these SIAM results identified potential SIEs according to an estimation of their interconnectedness within business networks and their size given by NCF/sector revenue for more than 6,600 SIE candidates in association with 45 NCFs. Given the focus on short lists of SIEs, we used measures of size and interconnectedness to rank SIEs and provide short lists. The list of top 25 overall SIEs based on their interconnectedness were provided in the memorandum to CISA.[6] We also provided a longer list of SIEs (i.e., top 100, 250, 500) and NCF-specific lists of SIEs in an accompanying interactive Excel spreadsheet.

The results of this SIAM prototype prioritize the inclusion of Phase 1 and 2 NCFs while including additional results for Phases 3–4, placing specific emphasis on Information Technology, Communication, Energy, Finance, Transportation, and Public Health sectors. One challenge for NCF-specific lists of SIEs is how many entities to include. We have provided a ranked list of the top 500 most systemically important entities, a ranked list of entities within each of the NCFs, and narrowed lists of the top 20 entities in each NCF. We arrived at lists of 20 after inspecting the distribution of systemic importance measures within each NCF. Although SIEs associated with the defense industrial base (DIB) were mentioned by the NRMC, the methods we used to map most NCFs at the time of this memorandum are not suitable for mapping the DIB. Although most NCFs can be mapped to one or several sectors, DoD contracted with firms in more than 95 percent of total six-digit NAICS codes from FY 2010 to FY 2018. We identified a possible path forward for identifying DoD contractors, which will be one focus of year two of this project.

Methods and Limitations of the SIAM Prototype

In support of the NRMC, HSOAC developed a methodology and prototype tool (1) to identify long lists of potential SIEs that directly support one or many NCFs and (2) that prioritizes potential SIE candidates based on their size and interconnectedness. We estimated an entity's size based on its sector revenue associated with a given NCF and estimated an entity's interconnectedness based on calculations of its network centrality within a large interfirm network of business relationships and supply chain linkages, a calculation that estimates the potential for systemic (upstream and downstream) impacts following idiosyncratic impacts to individual entities.[7] Importantly, sector revenue is a local measure which estimates an entity's systemic importance to an individual NCF, while network centrality is a global measure which estimates an entity's systemic importance within the entire interfirm network of connections and the global economy.

[6] In this list, we removed subsidiaries of parent companies.

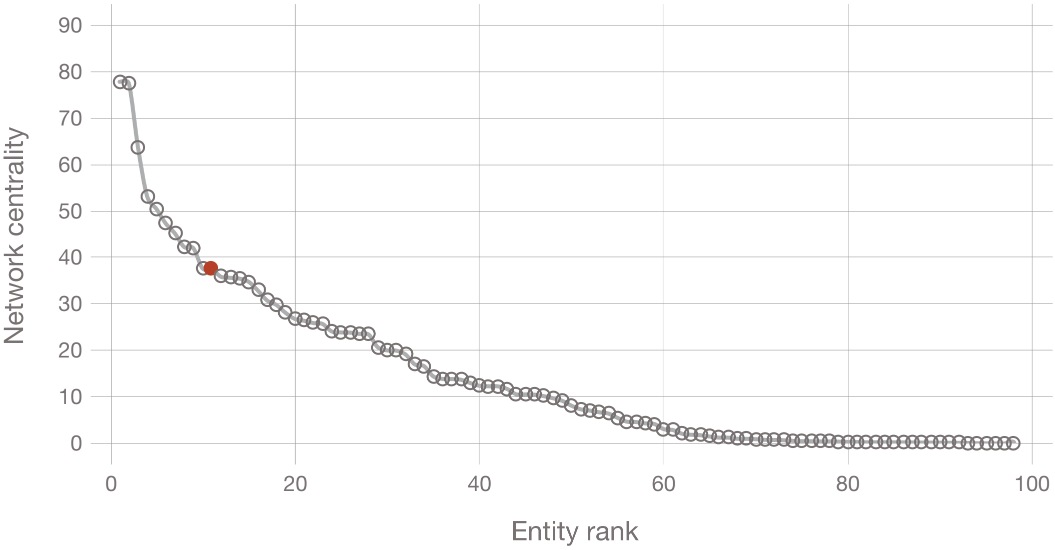
[7] For methodology on estimating the systemic risk posed by individual firms across the economy, see Welburn et al., 2020. The analysis leverages FactSet Revere for large data sets on business classifications and interfirm linkages. For full description, see FactSet, "The Basics," webpage, undated a; and FactSet, "FactSet Supply Chain Relationships," webpage, undated b, respectively.

The accompanying Excel file displayed the results of version 2.0 of the RAND SIAM Prototype. The interactive spreadsheet displayed each entity's name, its relative rank by network centrality, its network centrality score, the percentage of the entity's revenue from the NCF, its total sector sales, its overall entity sales, and the date of sales revenue. Also included in the spreadsheet was the entity type, which allows the user to easily identify public companies, subsidiaries, governments, holding companies, and other entity structures. Furthermore, in SIAM version 2.0, we narrowed lists of entities to only those which are U.S. domiciled and have at least 25 percent of their revenue in the NCF.[8]

Additionally, in the spreadsheet we provided a tool for analysts to address a question raised earlier in this memo; how many entities should be included in SIE lists? Figure D.1 displays the distribution of network centrality scores for the Provide Medical Care NCF. In the top 100 entities displayed in Figure D.1, there is significant variance in network centrality. As the rank of entity's increase along the graph, the value of network centrality decreases

FIGURE D.1

Finding the Knee-in-the-Curve for the Provide Medical Care National Critical Function



NOTE: In this figure, the knee in the curve is identified by graphing entities to the network centrality measures. Although 12, identified by the red dot, is the natural break point in this figure, note that some NCFs have multiple possible break points.

[8] Including firms with low levels of revenue in each NCF might result in false positives in which large companies that are highly interconnected are a leader within the NCF. In this case, the entity would appear as a SIE for the NCF with only a small portion of revenue in that function. We used a 25-percent revenue floor as a natural break point to determine whether the entity is interconnected with the function. The source data for sector revenue is not bound between 0 and 100 percent. An entity might reflect a negative sector revenue if the entity had losses in that sector or show revenue greater than 100 percent if the entity had losses in a different sector.

sharply before gradually decreasing at smaller and smaller increments. We recommend identifying the knee-in-the-curve as a data-driven guide for answering the question of how many entities to include. In Figure D.1, that knee-in-the-curve is at roughly 12 firms, as identified by the single red dot on the curve. Although the knee-in-the-curve varies across NCFs, we have found that it is typically less than 20 firms after applying the aforementioned 25 percent sector revenue threshold. Consequently, we provided lists of the top 20 firms in each NCF in the Excel tool.

However, the SIE prototype did reflect ongoing work with limitations and caveats. Results presented in the memorandum provided to the NRMC and the associated Excel spreadsheet were preliminary results of RAND-HSOAC research and, although informally reviewed, had not been formally reviewed or edited. Furthermore, although the current methodology is based on estimations of size and interconnectedness, additional work is required to estimate substitutability and, similarly, dynamic impacts following idiosyncratic disruption. Additionally, this analysis covered 45 of the 55 NCFs and, through measures such as revenue and network centrality, had a largely economic focus. SIEs in the 10 remaining NCFs and the inclusion of measures that emphasize national security, public health and safety, and marginalized populations will be included in the future. Additionally, the entities covered in the delivered SIAM prototype were biased toward corporations and, particularly, corporations that are publicly listed because of higher fidelity data. Consequently, the results might have underestimated (or entirely missed) the systemic importance of some public utilities.

The results, methods, and limitations presented in this appendix reflect the preliminary SIAM results delivered to the NRMC in February 2022. For a complete description of the methodology and limitations of the current SIAM model, please see Chapter Three.

The tasking to help respond to rapidly changing guidance and world events is not intended as a long-term role for HSOAC, but to backfill the understaffed and frequently tasked analysis function for the NRMC. Long term, the multiple aspects of the HSOAC SIE support are intended to provide the analytic function with the expertise, tools, and data sets it will need to continue to respond to real-time requests for information, while continuing to support national strategic risk analysis and planning.

Abbreviations

API	application programming interface
CII	Core Infrastructure Initiative
CISA	Cybersecurity and Infrastructure Security Agency
CITPS	Critical IT Products and Services
COVID-19	coronavirus disease 2019
CVSS	Common Vulnerability Scoring System
DHS	U.S. Department of Homeland Security
DoD	U.S. Department of Defense
EO	Executive Order
EPSS	Exploit Prediction Scoring System
FIRST	Forum of Incident Response and Security Teams
FSOC	Financial Stability Oversight Council
G-SIB	global systemically important bank
HSOAC	Homeland Security Operational Analysis Center
IT	information technology
NAICS	North American Industry Classification System
NCF	national critical function
NIST	National Institute of Standards and Technology
NRMC	National Risk Management Center
OCE	Office of the Chief Economist
RBICS	Revere Business Industry Classification System
SBOM	software bill of materials
SIAM	Systemic Importance Analytic Model
SICI	systemically important critical infrastructure

SIE	systemically important entity
SIFI	systemically important financial institution
USAGE	USA General Equilibrium

References

Acharya, V. V., and M. Richardson, "Implications of the Dodd-Frank Act," *Annual Review of Financial Economics*, Vol. 4, No. 1, 2012, pp. 1–38.

Adaptation Clearinghouse, "Social Vulnerability Index (SoVI)," webpage, undated. As of April 12, 2022:
https://www.adaptationclearinghouse.org/resources/social-vulnerability-index-sovi.html

Agency for Toxic Substances and Disease Registry, "CDC/ATSDR Social Vulnerability Index," Centers for Disease Control and Prevention, 2022. As of April 12, 2022:
https://www.atsdr.cdc.gov/placeandhealth/svi/index.html

Bank for International Settlements, "Global Systemically Important Banks: Assessment Methodology and the Additional Loss Absorbency Requirement," webpage, November 23, 2021. As of April 8, 2022:
https://www.bis.org/bcbs/gsib/.

Bank for International Settlements, "The G-SIB Assessment Methodology—Score Calculation," Basel Committee on Banking Supervision, 2014.

Bisias, Dimitrios, Mark Flood, Andrew W. Lo, and Stavros Valavanis, "A Survey of Systemic Risk Analytics," Office of Financial Research, working paper, No. 0001, 2012. As of June 16, 2022:
https://www.financialresearch.gov/working-papers/files/OFRwp0001_BisiasFloodLoValavanis_ASurveyOfSystemicRiskAnalytics.pdf

BitSight, homepage, undated. As of April 12, 2022:
https://www.bitsight.com/

Bodeau, Deborah J., and Catherine D. McCollum, *System-of-Systems Threat Model*, McLean, Va.: Homeland Security Systems Engineering and Development Institute, 2018.

Bunge, Jacob, "JBS Paid $11 Million to Resolve Ransomware Attack," *Wall Street Journal*, June 9, 2021. As of April 7, 2022:
https://www.wsj.com/articles/jbs-paid-11-million-to-resolve-ransomware-attack-11623280781.

CISA—*See* Cybersecurity and Infrastructure Security Agency.

Clancy, Noreen, Melissa L. Finucane, Jordan R. Fischbach, David G. Groves, Debra Knopman, Karishma V. Patel, and Lloyd Dixon, *The Building Resilient Infrastructure and Communities Mitigation Grant Program: Incorporating Hazard Risk and Social Equity into Decisionmaking Processes*, Homeland Security Operational Analysis Center operated by the RAND Corporation, RR-A1258-1, 2022. As of April 12, 2022:
https://www.rand.org/pubs/research_reports/RRA1258-1.html

CVE, "Overview," webpage, undated. As of April 12, 2022:
https://www.cve.org/About/Overview

Cybersecurity and Infrastructure Security Agency, "CISA Strategic Intent," webpage, undated a. As of April 12, 2022:
https://www.cisa.gov/publication/strategic-intent

Cybersecurity and Infrastructure Security Agency, "National Infrastructure Simulation and Analysis Center," webpage, undated b. As of April 12, 2022:
https://www.cisa.gov/NISAC

Cybersecurity and Infrastructure Security Agency, “National Risk Management,” webpage, undated c. As of April 12, 2022:
https://www.cisa.gov/national-risk-management

Cybersecurity and Infrastructure Security Agency, “Software Bill of Materials,” webpage, undated d. As of April 12, 2022:
https://www.cisa.gov/sbom

Cybersecurity and Infrastructure Security Agency, “National Critical Functions Set,” fact sheet, April 2019. As of April 8, 2022:
https://www.cisa.gov/sites/default/files/publications/national-critical-functions-set-508.pdf

Cybersecurity and Infrastructure Security Agency, *National Critical Functions: Status Update to the Critical Infrastructure Community*, Washington, D.C.: U.S. Department of Homeland Security, July 2020a. As of October 18, 2022:
https://www.cisa.gov/sites/default/files/publications/ncf-status-update-to-critical-infrastructure-community_508.pdf

Cybersecurity and Infrastructure Security Agency, *Resource Planning Guidance: FY2023–2027*, Washington, D.C.: U.S. Department of Homeland Security, December 2020b.

Cyentia Cybersecurity Research Library, “2020 State of the Software Supply Chain,” webpage, September 1, 2020. As of March 31, 2022:
https://library.cyentia.com/report/report_005128.html

Daly, Mary C., Bart Hobijn, and Joseph H. Pedtke, “Disappointing Facts About the Black-White Wage Gap,” FRBSF Economic Letter, Vol. 26, 2017, pp. 1–5.

Datanyze, “About,” webpage, undated. As of April 15, 2022:
https://www.datanyze.com/about

De Bandt, Olivier, and Philipp Hartmann, “Systemic Risk: A Survey,” European Central Bank, working paper, No. 35, November 2000.

DHS—*See* U.S. Department of Homeland Security.

European Central Bank, Financial Stability Review, European Central Bank, IV Special Features, 2010. As of May 26, 2022:
https://www.ecb.europa.eu/pub/pdf/fsr/art/ecb.fsrart201006_04.en.pdf

EO—*See* Executive Order.

Executive Order 13985, “Advancing Racial Equity and Support for Underserved Communities Through the Federal Government,” January 20, 2021.

Executive Order 14028, “Executive Order on Improving the Nation’s Cybersecurity,” May 12, 2021.

FactSet, “The Basics,” webpage, undated a. As of April 12, 2022:
https://insight.factset.com/resources/factset-revere-business-industry-classifications-datafeed

FactSet, “FactSet Supply Chain Relationships,” webpage, undated b. As of April 12, 2022:
https://www.factset.com/marketplace/catalog/product/factset-supply-chain-relationships

Federal Trade Commission, “Equifax Data Breach Settlement,” webpage, February 2022. As of April 7, 2022:
https://www.ftc.gov/enforcement/refunds/equifax-data-breach-settlement

Financial Stability Board, International Monetary Fund, and Bank for International Settlements, *Guidance to Assess the Systemic Importance of Financial Institutions, Markets, and Instruments: Initial Considerations—Background Paper*, October 28, 2009. As of May 27, 2022: https://www.fsb.org/wp-content/uploads/r_091107d.pdf

Finucane, Melissa L., Linnea Warren May, and Joan Chang, *A Scoping Literature Review on Indicators and Metrics for Assessing Racial Equity in Disaster Preparation, Response, and Recovery*, Santa Monica, Calif.: RAND Corporation, RR-A1083-1, 2021. As of March 29, 2022: https://www.rand.org/pubs/research_reports/RRA1083-1.html

FIRST, "Common Vulnerability Scoring System SIG," webpage, undated a. As of April 12, 2022: https://www.first.org/cvss/

FIRST, "Exploit Prediction Scoring System (EPSS)," webpage, undated b. As of April 12, 2022: https://www.first.org/epss/

Flavelle, Christopher, "Why Does Disaster Aid Often Favor White People?" *New York Times*, June 7, 2021. As of April 12, 2022: https://www.nytimes.com/2021/06/07/climate/FEMA-race-climate.html

Flowers, Catherine Coleman, "America's Dirty Shame: Living amid Raw Sewage," *Anglican Theological Review*, Vol. 100, No. 1, 2018, pp. 129–135.

Greenspan, Alan, *Remarks at a Research Conference on Risk Measurement and Systemic Risk, Statements and Speeches of Alan Greenspan*, Washington, D.C.: U.S. Federal Reserve, 1995. As of June 16, 2022: https://fraser.stlouisfed.org/title/statements-speeches-alan-greenspan-452/remarks-a-research-conference-risk-measurement-systemic-risk-washington-dc-8552/fulltext

Greenspan, Alan, "Never Saw it Coming: Why the Financial Crisis Took Economists by Surprise," *Foreign Affairs*, Vol. 92, No. 6, November/December 2013, pp. 88–96.

Greig, Jonathan, "Google Announces Scorecard V4 in Partnership with GitHub and OpenSSF," *ZDNet*, January 19, 2022. As of April 12, 2022: https://www.zdnet.com/article/google-announces-scorecard-v4-in-partnership-with-github-and-openssf/

Hallegatte, Stephane, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg, *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters*, Climate Change and Development, Washington, D.C.: World Bank, 2017.

Institutional Shareholder Services, homepage, undated. As of April 12, 2022: https://iss-cyber.com/

Karageorge, Eleni, "The Unexplainable, Growing Black-White Wage Gap," *Monthly Labor Review*, Vol. 140, November 2017.

Liu, Y., A. Sarabi, J. Zhang, P. Naghizadeh, M. Karir, M. Bailey, and M. Liu, "Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents," USENIX Security, Washington, D.C., August 2015.

McDermott, Melanie, Sango Mahanty, and Kate Schreckenberg, "Examining Equity: A Multidimensional Framework for Assessing Equity in Payments for Ecosystem Services," *Environmental Science and Policy*, Vol. 33, 2013, pp. 416–427.

Miro, Michelle E., Andrew Lauland, Rahim Ali, Edward W. Chan, Richard H. Donohue, Liisa Ecola, Timothy R. Gulden, Liam Regan, Karen M. Sudkamp, Tobias Sytsma, Michael T. Wilson, and Chandler Sachs, *Assessing Risk to the National Critical Functions as a Result of Climate Change*, Homeland Security Operational Analysis Center operated by the RAND Corporation, RR-A1645-7, 2022. As of April 5, 2022:
https://www.rand.org/pubs/research_reports/RRA1645-7.html

Montgomery, Mark, and Trevor Logan, "Poor Cybersecurity Makes Water a Weak Link in Critical Infrastructure," Foundation for Defense of Democracies, 2021. As of April 12, 2022:
https://www.fdd.org/analysis/2021/11/18/poor-cybersecurity-makes-water-a-weak-link-in-critical-infrastructure/

Mueller, J. Tom, and Stephen Gasteyer, "The Widespread and Unjust Drinking Water and Clean Water Crisis in the United States," *Nature Communications*, Vol. 12, No. 1, 2021, pp. 1–8.

Nagle, Frank, Jessica Wilkerson, James Dana, and Jennifer L. Hoffman, *Vulnerabilities in the Core: Preliminary Report and Census II of Open Source Software*, Linux Foundation Core Infrastructure Initiative, Linux Foundation, February 2020.

Nash, Kim S., Sara Castellanos, and Adam Janofsky, "One Year After NotPetya Cyberattack, Firms Wrestle With Recovery Costs," *Wall Street Journal*, June 27, 2018. As of April 14, 2022:
https://www.wsj.com/articles/one-year-after-notpetya-companies-still-wrestle-with-financial-impacts-1530095906

National Institute of Standards and Technology, "Definition of Critical Software Under Executive Order (EO) 14028," October 13, 2021. As of April 12, 2022:
https://www.nist.gov/system/files/documents/2021/10/13/EO%20Critical%20FINAL.pdf

Nolle, Daniel E., "U.S. Domestic and International Financial Reform Policy: Are G20 Commitments and the Dodd-Frank Act in Sync?" Board of Governors of the Federal Reserve System, International Finance Discussion Papers, No. 1024, July 2011. As of May 27, 2022:
https://www.federalreserve.gov/pubs/ifdp/2011/1024/ifdp1024.htm

Norton-Taylor, Richard, "Titan Rain: How Chinese Hackers Targeted Whitehall," *The Guardian*, Vol. 4, 2007.

Office of Personnel Management, "Cybersecurity Resource Center: Cybersecurity Incidents," webpage, undated. As of April 7, 2022:
https://www.opm.gov/cybersecurity/cybersecurity-incidents/

Okeowo, Alexis, "The Heavy Toll of the Black Belt's Wastewater Crisis," *New Yorker*, November 23, 2020. As of May 26, 2022:
https://www.newyorker.com/magazine/2020/11/30/the-heavy-toll-of-the-black-belts-wastewater-crisis

Perlroth, Nicole, and David E. Sanger, "Cyberattacks Put Russian Fingers on the Switch at Power Plants, U.S. Says," *New York Times*, March 15, 2018. As of April 14, 2022:
https://www.nytimes.com/2018/03/15/us/politics/russia-cyberattacks.html

Peterson, Dana M., and Catherine L. Mann, *Closing the Racial Inequality Gaps: The Economic Cost of Black Inequality in the U.S.*, Citi GPS: Global Perspectives and Solutions, September 2020. As of April 14, 2022:
https://ir.citi.com/NvIUklHPilz14Hwd3oxqZBLMn1_XPqo5FrxsZD0x6hhil84ZxaxEuJUWmak51UHvYk75VKeHCMI%3D

Public Law 111-203, Dodd-Frank Wall Street Reform and Consumer Protection Act, 2010.

Public Law 115-232, John S. McCain National Defense Authorization Act for Fiscal Year 2019, 2018.

RiskRecon, homepage, undated. As of April 12, 2022:
https://www.riskrecon.com/

Robles, Frances, and Nicole Perlroth, "'Dangerous Stuff': Hackers Tried to Poison Water Supply of Florida Town," *The New York Times*, February 8, 2021. As of April 7, 2022:
https://www.nytimes.com/2021/02/08/us/oldsmar-florida-water-supply-hack.html

Romanosky, Sasha, John Bordeaux, Michael J. D. Vermeer, Jonathan W. Welburn, Aaron Strong, and Zev Winkelman, *Identifying Critical IT Products and Services*, Homeland Security Operational Analysis Center operated by the RAND Corporation, RR-A923-2, 2022. As of October 18, 2022:
https://www.rand.org/pubs/research_reports/RRA923-2.html

Ruhnau, Britta, "Eigenvector-Centrality—a Node-Centrality?" *Social Networks*, Vol. 22, No. 4, 2000, pp. 357–365.

Sawers, Paul, "Google and Microsoft Back the Alpha-Omega Project to Bolster Software Supply Chain," *VentureBeat*, February 1, 2022. As of April 12, 2022:
https://venturebeat.com/2022/02/01/google-and-microsoft-back-the-alpha-omega-project-to-bolster-software-supply-chain/

Schwarcz, Steven L., "Systemic Risk," *Georgetown Law Journal*, Vol. 97, No. 1, 2008, pp. 193–250. As of June 16, 2022:
https://heinonline.org/HOL/P?h=hein.journals/glj97&i=195

Schweizer, Pia-Johanna, "Systemic Risks—Concepts and Challenges for Risk Governance," *Journal of Risk Research*, Vol. 24, No. 1, 2019, pp. 78–93.

Security Scorecard, homepage, undated. As of April 12, 2022:
https://securityscorecard.com/

StatCounter, "Browser Market Share Worldwide: May 2021," webpage, undated. As of April 12, 2022:
https://gs.statcounter.com/browser-%20market-share#monthly-202105-202105-bar

Temple-Raston, Dina, "A 'Worst Nightmare' Cyberattack: The Untold Story Of The SolarWinds Hack," *NPR*, April 16, 2021. As of April 12, 2022:
https://www.npr.org/2021/04/16/985439655/a-worst-nightmare-cyberattack-the-untold-story-of-the-solarwinds-hack

Thornburgh, Nathan, "The Invasion of the Chinese Cyberspies," *Time*, Washington, D.C., August 29, 2005.

Torres-Arias, Santiago, "What is Log4j? A Cybersecurity Expert Explains the Latest Internet Vulnerability, How Bad It Is and What's At Stake," *The Conversation*, December 21, 2021. As of April 12, 2022:
https://theconversation.com/what-is-log4j-a-cybersecurity-expert-explains-the-latest-internet-vulnerability-how-bad-it-is-and-whats-at-stake-173896

Turner, Ani, "The Business Case for Racial Equity," *National Civic Review*, Vol. 105, No. 1, April 7, 2016, pp. 21–29.

Turton, William, and Kartikay Mehrotra, "Hackers Breached Colonial Pipeline Using Compromised Password," *Bloomberg*, June 4, 2021. As of April 7, 2022:
https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password

U.S. Code Title 42, Section 5195c, Critical Infrastructures Protections Act of 2001.

U.S. Cyberspace Solarium Commission, *Cyberspace Solarium Commission Report*, Washington, D.C., 2020. As of April 7, 2022:
https://www.solarium.gov/report

U.S. Department of Homeland Security, *DHS Risk Lexicon: 2010 Edition*, September 2010. As of March 31, 2022:
https://www.cisa.gov/sites/default/files/publications/dhs-risk-lexicon-2010_0.pdf

U.S. Department of Homeland Security, "DHS Announces Changes to Individual Assistance Policies to Advance Equity for Disaster Survivors," press release, September 2, 2021. As of April 12, 2022:
https://www.dhs.gov/news/2021/09/02/dhs-announces-changes-individual-assistance-policies-advance-equity-disaster

U.S. House of Representatives, "Securing Systemically Important Critical Infrastructure Act," Bill 5491, 117th Congress, October 5, 2021.

U.S. Water Alliance, "Closing the Water Access Gap in the United States: A National Action Plan," 2019. As of May 26, 2022:
http://uswateralliance.org/sites/uswateralliance.org/files/publications/Closing%20the%20Water%20Access%20Gap%20in%20the%20United%20States_DIGITAL.pdf

Vittorio, Andrea, "Merck's $1.4 Billion Insurance Win Splits Cyber From 'Act of War,'" *Bloomberg Law*, January 19, 2022. As of April 14, 2022:
https://news.bloomberglaw.com/privacy-and-data-security/mercks-1-4-billion-insurance-win-splits-cyber-from-act-of-war

Welburn, Jonathan W., and Aaron Strong, "Systemic Cyber Risk and Aggregate Impacts," *Risk Analysis*, 2021.

Welburn, Jonathan W., Aaron Strong, Florentine Eloundou Nekoul, Justin Grana, Krystyna Marcinek, Osonde A. Osoba, Nirabh Koirala, and Claude Messan Setodji, *Systemic Risk: It's Not Just in the Financial Sector*, Santa Monica, Calif.: RAND Corporation, RB-10112-RC, 2020. As of March 30, 2022:
https://www.rand.org/pubs/research_briefs/RB10112.html

Winkler, Inga T., and Catherine Coleman Flowers, "'America's Dirty Secret': The Human Right to Sanitation in Alabama's Black Belt," *Columbia Human Rights Law Review*, Vol. 49, 2017, pp. 181–228.

World Economic Forum, "Understanding Systemic Cyber Risk," Global Agenda Council on Risk and Resilience, white paper, October 2016.

World Economic Forum, "Beneath the Surface: Technology-Driven Systemic Risks and the Continued Need for Innovation," Future of Financial Services Series, October 28, 2021. As of May 26, 2022:
https://www.weforum.org/reports/beneath-the-surface-technology-driven-systemic-risks-and-the-continued-need-for-innovation

Young, Shalanda, Brenda Mallory, and Gina McCarthy, "The Path to Achieving Justice40," *The White House Briefing Room Blog*, July 20, 2021. As of April 12, 2022:
https://www.whitehouse.gov/omb/briefing-room/2021/07/20/the-path-to-achi

In response to the mounting specter of systemic cyber risks, the Cyberspace Solarium Commission recommended that Congress codify the concept of Systemically Important Critical Infrastructure—later renamed Systemically Important Entities (SIEs)—and that the Cybersecurity and Infrastructure Security Agency (CISA) be resourced to identify SIEs and support in the mitigation of their risks to support a broader national strategy of layered deterrence. In support of the CISA National Risk Management Center (NRMC), this report clarifies the concepts of SIEs and introduces a data-driven methodology for their identification and prioritization. Specifically, the authors identify SIEs by their potential to affect national critical functions (NCFs) and prioritize SIEs by measures of their size and interconnectedness.

This report builds on existing work regarding Critical IT Products and Services and extending the researchers' analysis to federal agencies and firms that install potentially vulnerable software, in addition to firms that write software. This report further documents systemic risks and cyber risks in software supply chains, past and ongoing analytical support to CISA, and current limitations, and it outlines a path for future work.

$28.00

ISBN-10 197740984-9
ISBN-13 9781977409843

52800

9 781977 409843

Cover image: Scanrail/Getty Images
RR-A1512-1