MITRE | SOLVING PROBLEMS FOR A SAFER WORLD®

DIVERSITY IN THE CYBER WORKFORCE: ADDRESSING THE DATA GAP

Dr. Irving Lachow

 PRS21-01226-4 1-28-2022

Table of Contents

Tables

Introduction

This report is the result of a grant awarded to The MITRE Corporation by the Hewlett Foundation to examine the challenges associated with producing a demographic baseline of the nation's cyber workforce. MITRE subsequently partnered with Aspen Digital—a program of the Aspen Institute—to carry out the study based on a literature review, workshops, and expert interviews. A draft copy of this report was circulated to all workshop participants and interview subjects to elicit substantive feedback. While we have done our best to incorporate the information and suggestions provided by all reviewers, the author is ultimately responsible for the content herein.

Section 1 summarizes the benefits of having a diverse cyber workforce and describes the current state of knowledge regarding the diversity of that workforce. **Section 2** summarizes the key findings of our study. It identifies the most significant challenges associated with the collection, analysis, and distribution of cyber workforce diversity data and provides insights based on our research and inputs from key stakeholders. **Section 3** provides actionable recommendations that address the challenges and leverage the insights described in **Section 2**.

Why Diversity Matters in Cyber

Cyber practitioners and national security officials often express alarm at the absence of diversity among cyber specialists in both the public and private sectors.[1] The global economy has an estimated 3.1 million unfilled cybersecurity positions. U.S.-based businesses and government agencies face a shortage of between 350,000 and 600,000 cybersecurity professionals and 56% of companies believe that their staffing shortfalls put them at moderate or extreme risk.[2]

Increasing the diversity, equity, and inclusion (DEI) of the cyber workforce can help address this workforce shortage while simultaneously having a positive impact on business growth and performance.[3] For example, companies whose executive teams are in the top quartile for gender and ethnic diversity are 21% and 33% more profitable, respectively, than similar companies with less diverse leadership. This is due to improved problem solving and idea generation:

[1]For example, see M. Miller, "Biden administration establishes program to recruit tech professionals to serve in government," *The Hill*, 30-Aug-2021. [Online]. Available: https://thehill.com/policy/cybersecurity/570068-biden-administration-establishes-program-to-recruit-tech-professionals.

[2]Data taken from International Information System Security Certificate Consortium (ISC2), "Cybersecurity Professionals Stand Up to a Pandemic - (ISC)2 Cybersecurity Workforce Study 2020," *www.isc2.org*, 2020. [Online]. Available: https://www.isc2.org/-/media/ISC2/Research/2020/Workforce-Study/ISC2ResearchDrivenWhitepaperFINAL.ashx; and Cyberseek.org, "*Cybersecurity supply and demand heat map*," Cybersecurity Supply and Demand Heat Map, 2022. [Online]. Available: http://www.cyberseek.org/heatmap.html.

[3]See V. Hunt, L. Yee, S. Prince, and S. Dixon-Fyle, "Delivering through diversity," *McKinsey & Company*, 18-Jan-2018. [Online]. Available: https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/delivering-through-diversity; K. Phillips, "How Diversity Makes Us Smarter," *Scientific American*, vol. 30, no. 3, Jul. 2021; D. Rock and H. Grant, "Why diverse teams are smarter," *Harvard Business Review*, 04-Nov-2016. [Online]. Available: https://hbr.org/2016/11/why-diverse-teams-are-smarter; and S. Lyons, "Council Post: The Benefits of Creating A Diverse Workforce," *Forbes*, 09-Sep-2019. [Online]. Available: https://www.forbes.com/sites/forbescoachescouncil/2019/09/09/the-benefits-of-creating-a-diverse-workforce/?sh=15cbc1d9140b.

"Diverse teams have been shown to be more likely to radically innovate and anticipate shifts in consumer needs and consumption patterns—helping their companies to gain a competitive edge."[4]

These benefits carry over to the security realm.[5] A diverse workforce can contribute to a better understanding of user behavior and the ever-evolving threat landscape. Additionally, diverse representation can help organizations identify and address implicit biases that may be impacting their ability to hire and retain talent, develop new products and services, and understand market demand across a broader demographic base.[6]

Addressing the nation's growing cybersecurity needs will require a concerted effort that includes people from across the full spectrum of our nation's citizenry. Recognizing this, multiple initiatives spearheaded by a mix of nonprofits, companies, and government agencies are trying to change how the cyber community engages, trains, hires, and retains employees to foster a more diverse and inclusive workforce. Notable examples include #ShareTheMicInCyber, the Gula Tech Foundation, the Aspen Institute's Cyber Workforce Coalition, and organizations such as Women in Cybersecurity and Cyversity. More recently, over 30 Chief Executive Officers from industry, academia, and civil society signed on to key findings of the Aspen Institute's Action to Catalyze (ACT) report, which calls on industry to transform how it brings underrepresented talent into the tech industry.[7] These efforts are incredibly important, but without having a solid baseline of data on the state of the cybersecurity workforce, it is difficult to assess the impact that these initiatives are having. This makes it challenging to determine which efforts are deserving of additional support and which may need to be modified in some way.

One final observation: The U.S. government (USG) also needs to develop a clearer picture of the diversity of its cyber workforce—a key point made by Cyberspace Solarium Commission.[8] First, this will help the USG develop a larger and more effective cyber cadre in a critical area of concern for the nation. Second, it will set a good example for the rest of the country. Third, the USG is both a consumer of, and a provider of, cyber talent. It is important for the USG to remain in step with industry trends and perhaps even to drive them in desired directions.

[4]S. Dixon-Fyle, K. Dolan, V. Hunt, and S. Prince, "Diversity wins: How inclusion matters," *McKinsey & Company*, May-2020. [Online]. Available: https://www.mckinsey.com/featured-insights/diversity-and-inclusion/diversity-wins-how-inclusion-matters.

[5]S. John, "Why we need more diversity in cybersecurity," *Microsoft News Centre Europe*, 28-May-2020. [Online]. Available: https://news.microsoft.com/europe/features/why-we-need-more-diversity-in-cybersecurity/.

[6]C. Stewart, "Systemic Racism Is a Cybersecurity Threat," *Council on Foreign Relations*, 16-Jun-2020. [Online]. Available: https://www.cfr.org/blog/systemic-racism-cybersecurity-threat.

[7]Catalyze Tech Working Group, "The ACT Report: Action to Catalyze Tech, A Paradigm Shift for DEI," Published by the Aspen Institute and Snap Inc. Oct-2021. [Online]. Available: ACTReport.com.

[8]Bate, Laura, "Cyberspace Solarium Commission - Workforce White Paper", White Paper, Sep-2020. [Online]. Available: https://www.solarium.gov/public-communications/workforce-white-paper.

Data as a Diversity Accelerant

A more rigorous, coordinated approach to DEI depends in part on a baseline understanding of the current demographic composition of the cyber field. The workshops and interviews performed for this study revealed near-universal agreement that **producing and sharing more comprehensive data on the makeup of the cyber workforce would support and accelerate operational changes to education, recruitment, training, and retention practices**. Identified benefits include:

- Providing a baseline of cyber workforce DEI for assessing the effectiveness of policies and programs. This can occur both at the enterprise level (e.g., evaluating whether a recruitment strategy increases the number of women who apply to security roles) and at a regional or national level (e.g., evaluating a K-12 program aiming to attract underrepresented groups to the cybersecurity field).
- Identifying priority areas for additional investment aimed at improving the diversity of the cyber workforce
- Enabling organizations and industry sectors to benchmark themselves in relation to similar entities. By leveraging the spirit of competition among rivals, this can help create marketing, branding, and recruitment incentives that drive organizations toward greater diversity without legal mandates.
- Highlighting the geographical roots of disparities within and across national and multinational organizations, which may be important in understanding the roles of different laws, policies, and cultures.
- Providing greater insight into issues like pay equity and promotion rates.
- Helping workers make informed decisions about potential employers by showing how seriously those employers take their public commitments to diversity and inclusion.

Where Is the Data Today?

To understand whether current data collection efforts might meet the cyber community's needs for greater visibility into the makeup of the cyber workforce, our team examined industry studies, government data, and corporate reporting related to DEI, equal employment opportunities, and social responsibility. All available data sources have shortcomings, making it difficult to understand and track the changing makeup of the cyber workforce in a comprehensive and consistent way.

Generally, studies that focus on cyber workforce do not focus on diversity. Those studies or reports that gather extensive diversity data focus on broad workforce categories like "information technology" (IT) or "science, technology, engineering, and math (STEM)." The Information System Security Certification Consortium, also known as (ISC)2, has published two reports, one in 2018 and another in 2020, that specifically address the diversity of the cyber workforce. These studies provide the most useful data on this topic and are often cited by the media. As helpful as they are in providing a basic picture of the field, the (ISC)2 reports do not break down the overall cyber workforce into different job functions. Without this granularity, it is hard to know where the survey drew the line as to what constitutes the cybersecurity workforce. That makes it hard to compare between different studies or track progress over time.

Government sources like the Bureau of Labor Statistic (BLS) and National Science Foundation (NSF) are limited in other ways. The BLS data looks at only one role within the cybersecurity field: information security analyst. NSF data has focused on the STEM workforce broadly rather than the cyber field.

Corporate reporting sheds little light on the state of DEI in the cyber workforce. Our team examined public documents from 10 companies and found

that most focused on their entire workforce with no specific breakouts for the information technology field, let alone for cyber positions.[9] These reports, summarized in **Table 1**, provide varying levels of information on workforce diversity, and none focus on the cyber workforce specifically. In addition, few reports provide data related to disability, veteran status, sexual orientation, gender identity, age, and education.

Table 2 summarizes the tradeoffs inherent in existing data sources. A more detailed examination of the best available sources of cyber workforce data is presented in **Table 3**.

In addition to examining the cyber workforce directly, we looked for lessons learned from other industries. One such field is artificial intelligence (AI). Findings from a recent report co-authored by MITRE and the Center for Security and Emerging

COMPANY NAME		REPORT TYPE	YEAR	EMPLOYEE BASE		DIVERSITY CATEGORIES								PROFESSIONAL TITLE AND ROLES				
				US	Global	Sex	Race & Ethnicity	Disability	Veteran Status	Sexual Orientation	Gender Identity	Age	Education	Individual Contributor	Managers	Executive	Technical Roles	New Hires
Tech	Adobe	Diversity & Inclusion Report	2020	x	x	x	x							x	x	x	x	x
	Dell	Diversity & Inclusion Report	2020	x	x	x	x								x		x	
	IBM	Diversity & Inclusion Report	2020	x	x	x	x								x	x	x	x
	Oracle	EEO Employer Information Report	2018			x	x											
	Snap, Inc	EEO Employer Information Report	2020	x	x	x	x	x	x	x	x	x	x			x	x	x
Non-Tech	Coca-Cola	Diversity Annual Report	2020	x	x	x	x											
	Nike, Inc.	Business & Environmental, Social	FY20	x	x	x	x											
	Target	and Governance Report	2020	x	x	x	x											
	Walmart	Impact Report	2020	x	x	x	x											x
	Wells-Fargo	Corporate Responsibility Report	2020	x	x	x	x	x	x			x						x

TABLE 1. SAMPLE OF CORPORATE DIVERSITY REPORTS

SOURCE	DIVERSITY SCOPE	CYBER RELEVANCE
Government Data	🟢	🔴
Corporate DEI Reports	🟡	🔴
Industry Studies	🟡	🟢

TABLE 2. COMPARISON OF CYBER WORKFORCE DIVERSITY DATA SOURCES

[9]The DEI reports that we examined came from: Adobe, Coca-Cola, Dell, IBM, Nike, Oracle, Snap, Target, Walmart, and Wells Fargo.

ORGANIZATION OR AGENCY	SOURCE	YEAR	WORKFORCE SCOPE	DIVERSITY CATEGORIES/TRAITS
Bureau of Labor Statistics (BLS)	Current Population Surveys	2020	Information Security Analyst	• Sex • Race & Ethnicity • Age • Disability
Information System Security Certification Consortium and Frost & Sullivan	Innovation Through Inclusion: The Multicultural Cybersecurity Workforce	2018	Cybersecurity	• Sex • Age • Race & Ethnicity
Information System Security Certification Consortium	(ISC)2 Cybersecurity Workforce Study: Cybersecurity Professionals Stand up to a Pandemic	2020	Cybersecurity (defined as information technology professionals who spend more than 25% of their time engaged in cybersecurity tasks)	• Sex • Age
McKinsey & Company	Diversity Wins – How Inclusion Matters	2020	General workforce	• Sex • Race & Ethnicity
National Center for Science and Engineering Statistics (NCSES)	Women, Minorities, and Persons with Disabilities in Science and Engineering	2021	Science and Engineering	• Sex • Disability • Race & Ethnicity • Age

TABLE 3. SELECTED SOURCES OF CYBER WORKFORCE DATA

Technology at Georgetown University echo the challenges found in the cyber field. Among the key observations from that report:

- There is an inconsistent approach to identifying AI talent.
- AI talent can be divided into technical and non-technical aspects.
- Technical talent can be further categorized into those with directly applicable traits (e.g., a data science degree) and those who can do AI work but may not be identified as having that capability (e.g., electrical engineers). The latter, who are viewed as "AI-adjacent," could perform technical AI functions with minimal training.
- Non-technical talent includes "those in roles that complement technical talent, including acquisition personnel and program managers."[10]

These observations are directly applicable to the cyber workforce. Like the AI workforce, the cyber field includes both technical and non-technical talent. Within the technical talent category, one can also identify those that are "cyber-adjacent" in terms of skills (e.g., a scientist who learned to code while doing data analyses for their research). The AI report, like many of the cyber studies in the literature, recommends stronger efforts to identify and track talent. Like the cyber field, the AI field lacks detailed data on the diversity in its workforce.

[10]D. Gehlhaus, R. Hodge, L. Koslosky, K. Goode, and J. Rotner, "The DOD's Hidden Artificial Intelligence Workforce," *Center for Security and Emerging Technology, Policy Brief*, Sep-2021, p. 4. [Online]. Available: https://cset.georgetown.edu/publication/the-dods-hidden-artificial-intelligence-workforce/.

Summary of Findings

Our research yielded the following insights:

Types of Data Needed

- Cyber workers are spread throughout an enterprise. While it is important to capture data on engineers and computer scientists, we also need to identify managers, lawyers, policymakers, and non-technicians who are key players in the cyber arena.
- It is unclear which aspects of diversity can and should be collected. Biological characteristics like age, race, and birth sex are the easiest to capture. A full exploration of diversity would need to include ethnicity, sexual orientation and gender identity, neurodiversity, and perhaps other aspects as well. However, the collection of such data would be difficult. This issue needs further exploration.

Data Collection Process

- Data collection should be voluntary.
- Data anonymization is critical.
- The security and privacy of the collected data must be paramount.
- Impacted communities need to be part of the data collection process and survey design. If they are not, data collection efforts might unintentionally perpetuate biases found in the field.
- A one-time snapshot of the cyber workforce is not helpful. To track progress and assess the utility of different policies and initiatives, it is imperative to gather longitudinal data, based on consistent definitions and criteria.

Analysis and Sharing

- Data analysis is critical: Decision makers rely on analytical findings to make decisions—raw data is not useful for their needs.
- Two sharing models deserve further exploration:

  Information Analysis and Sharing Organizations/Centers (ISAOs/ISACs)

  The Aviation Safety Information Analysis and Sharing (ASIAS) structure

Organization

- The data collection, analysis, and dissemination functions must be run by a single organization. There are several requirements for that organization:

  It must be trusted by those providing the data. There cannot be any real or perceived conflicts of interest.

  It must be able to safeguard the data.

  It must employ or have access to a team of organizational psychologists, economists and/or statisticians, and cyber experts.

  It must be able to withstand ebbs and flows in political sentiment around this issue.

- Not-for-profit organizations provide the best option for taking on the role of collecting, storing, and analyzing cybersecurity diversity data. Several types of not-for-profit organizations appears to satisfy the criteria needed to do the job, including, but not limited to, industry associations, information sharing and analysis organizations/centers, think tanks, federally funded research and development organizations, and university-affiliated research centers.

Funding

- Developing a useful demographic picture of the nation's cyber workforce will require information gathering, analysis, and dissemination over a period of years. This enterprise will require sustained funding that will likely cost several hundred thousand dollars per year.
- Industry is unlikely to fund such an activity initially and may not fund it at all. For this effort to succeed, it will require a consistent source of funds over several years. That type of funding is best provided by the U.S. government.

Analyses and Observations

Any attempt to address the current shortage of cyber workforce diversity data will need to determine (1) what kind of data to collect; (2) the process for gathering, processing, analyzing, and sharing the data; (3) the type of organization that should perform these tasks; and (4) who will pay for the effort. These questions are critical to understanding the feasibility and costs associated with options for improving our understanding of cyber workforce diversity. This section elaborates on the summary provided above as it delves into our exploration of the topic.

What Data Is Collected?

There are numerous obstacles associated with attempts to capture data about the diversity of the cyber workforce. The first centers on the inherent ambiguity of two key terms: diversity and workforce. As discussed earlier, diversity can be defined according to many criteria: gender, sexual orientation, race, ethnicity, age, nationality, geography, etc.

This definitional issue causes several challenges. First, as a practical matter, gathering information on many dimensions of diversity increases the complexity, cost, and time of data collection. In addition, individuals might be hesitant to disclose certain aspects of diversity information for fear that such information could be used as a pretext for adverse actions. For this reason, governments often expressly prohibit employers from requesting certain information from employees. Finally, data collection may be challenging for multi-national organizations that need to comply with a myriad of laws, regulations, and cultural norms.

Because different employers collect varying data on diversity, attempts to create a standardized dataset from existing sources will be complicated. If one company only gathers gender data while another focuses only on race, combining the two datasets will generate few useful insights. Drawing a comprehensive picture of diversity calls for new, standardized data collection processes.

Similarly, the many definitions of a "cybersecurity role" make it difficult to create a unified picture of the cyber workforce. Cybersecurity is a horizontal function that cuts across many job families. A technical employee whose job description omits a reference security (e.g., a network administrator) might in fact perform a central security role. Moreover, cyber roles encompass many non-technical functions that can fall under different business verticals. A human resources manager or attorney at a cyber incident response firm is supporting that organization's cyber mission even if their specific duties are not related to security operations. Corporate leaders like Chief Information Officers and Chief Financial Officers have tremendous influence on cybersecurity and may have little to no technical background in the area. On the government side, policymakers like the Deputy Secretary of Defense (Cyber Policy), the Director of the Cybersecurity

and Infrastructure Security Agency (CISA), and National Security Council staff may come to the cyber field with backgrounds in law or national security rather than engineering or computer science. Such individuals would not be counted as cybersecurity staff in some of the existing surveys, yet their roles are critical and understanding the diversity of people at that level is important.[11]

Designing a standardized data collection process that can account for these dynamics is difficult. The National Initiative for Cybersecurity Education Workforce Framework for Cybersecurity (a.k.a. the "NICE Framework") attempts to standardize the lexicon and job classifications associated with the cybersecurity field. For example, the NICE Framework identifies seven different categories, 33 specialty areas, and 52 different cyber roles.[12] It includes both technical and non-technical roles (examples of the latter include "Cyber Legal Advisor" and "Cyber Policy and Strategy Planner") and it is constantly being updated based on user feedback.

There is no question that the NICE Framework is useful. At the same time, our study found that many industry participants do not yet use the NICE Framework. Several companies also expressed concerns that the NICE Framework's focus on cybersecurity job functions misses the breadth of non-technical roles in the field.

In summary, **understanding the diversity of the cyber field requires one to capture the full range of cyber workers**. This broader focus requires an expanded aperture: data gathering efforts must go beyond narrow definitions of cybersecurity. For example, attempts to gather company-wide cyber data will need to move beyond the chief information officer and include human resources, and possibly organizations focused on research and development. Collecting longitudinal data is also important because such information can help organizations track which workers are leaving the cybersecurity workforce and why they are making those choices.

How Is Data Collected?

The collection of diversity data is a sensitive endeavor. Any initiative to collect comprehensive data on diversity from private employers will need to navigate an ocean of legal, procedural, and cultural obstacles. Unsurprisingly, workshop discussions and interviews with experts revealed that mandatory data collection—requiring private employers to record diversity statistics across their cyber workforce—is likely to be a political nonstarter and is not feasible as a practical or legal matter. The federal government would need to pass legislation or create policies that require this type of data disclosure. Doing that would require government leaders to work through a variety of thorny issues, such as which agency(ies) should be involved in the effort, which Congressional committees/subcommittees would have oversight, how civil liberties protections would apply to the collected data, how the activity would be funded, and what penalties would be imposed for noncompliance. An additional consideration is that a legal mandate requiring companies to share diversity data for cyber workers would provoke serious opposition, and potentially legal challenges, from powerful stakeholders.

[11]The The Biden Administration appointed the first non-white, non-male to the role of Deputy Assistant Secretary of Defense (Cyber Policy). See U.S. Department of Defense, "Mieke Eoyang, Deputy Assistant Secretary of Defense for Cyber Policy," *U.S. Department of Defense*. [Online]. Available: https://www.defense.gov/About/Biographies/Biography/Article/2505290/mieke-eoyang/.

[12]NICCS, "Workforce Framework for cybersecurity (NICE framework)," *National Initiative for Cybersecurity Careers and Studies*, 29-Jul-2021. [Online]. Available: https://niccs.cisa.gov/workforce-development/cyber-security-workforce-framework.

A voluntary data collection process appears less problematic than a mandate, though it does present its own unique challenges. The success of a voluntary effort depends on sustained and broad-based stakeholder engagement to recruit participants. Organizations that choose to participate in the activity may provide incomplete data. Staff departures could leave data collectors without a point of contact who understands the context and importance of the initiative. Finally, a voluntary process could suffer from data bias if some organizations cherry-pick the information they submit.

Despite this litany of challenges, we believe that **a voluntary approach is the best avenue for moving forward**. The key to success will be providing participants with sufficient value to motivate their involvement. At the same time, workshop participants made clear that two issues must be addressed if there is any hope for a data gathering effort to succeed: anonymity and security/privacy.

- If specific organizations providing data are identified by name, there is a risk that low-performing companies will avoid sharing data that hurts their brand. If high-performing companies do provide data—to garner positive publicity—then the dataset will be skewed and not representative of the state of cyber workforce diversity across the nation. There is also a possibility, specifically mentioned by several workshop participants, that even high-performing companies might resist sharing data because of fears that competitors will attempt to poach staff from their ranks.
- The solution to these concerns is to ensure that the data and analyses produced and shared by the collecting organization be anonymized. Anonymizing data would also address legal and ethical concerns. Survey participants could still generate positive publicity by highlighting their participation in the overall effort. This would demonstrate their commitment to DEI in the cyber workforce without revealing the details of the data they provided.
- Workshop participants made clear that they would be deeply concerned about the security of data they provided to a third party (be it government, a not-for-profit organization, or a private company). They want assurances that their data will be protected from misuse, leaks, and theft. The real issue here is trust. The organization that is collecting, storing, and analyzing the diversity information must demonstrate that it does not have any real or apparent conflicts of interest with those providing the data. It must also demonstrate that it follows industry leading practices in cybersecurity and privacy.
- The data gathering and analysis organization will need to communicate the objectives of the data collection activity clearly and transparently. It must also demonstrate the value proposition of the activity to potential participants.
- One final observation: To avoid perpetuating biases and discounting the views of underrepresented populations, it is critical that the data collection process include the perspectives of those who will be counted. There are a variety of approaches for ensuring that this takes place.[13] The data gathering organization will need to determine which method (or methods) is most applicable to its processes.

[13]For example, a set of tools is provided free of charge via MITRE's Social Justice Platform: https://sjp.mitre.org/.

Options for Data Sharing and Analysis

One of the key findings from our workshops, which was backed up by our discussions with the NSF's National Center for Science and Engineering Statistics, is that raw data alone has limited utility. **Policymakers and corporate leaders care about analytics that help them make evidence-based decisions**. Useful insights come from analyses that answer specific questions. The logical inference from this observation is that the organization that is responsible for gathering and storing/sharing cyber workforce diversity data will need to have the ability to conduct analyses that address the needs of key stakeholders. This implies the need to rely on an organization—be it a company, a government agency, or a not-for-profit—with access to data scientists, economists and/or statisticians, psychologists, and cybersecurity experts.

The issue of data sharing also came up during our workshops. Two suggestions were made. One was to examine data sharing models used by ISACs and ISAOs—particularly those that focus on the sharing of cyber threat information. These organizations generally work by enabling members to share information directly with each other. In this model, the ISAC/O operator serves primarily as a clearinghouse, though in some cases it also provides value-added services, such as issuing alerts and sharing best practices, that benefit all members.

The Federal Aviation Administration's ASIAS system provides an entirely different model.[14] In this case, both industry and government participants agree to provide their information to a trusted third party (TTP). The TTP pools everyone's data, analyzes the data, and shares its findings back out to the group. A member's data is never shared with other members—the TTP is the only entity that sees the raw data. What's more, the analyses that it conducts each year are determined by the members via a steering committee. Thus, members have a say in the operations of the TTP. Another advantage of this model is that it is designed to incorporate both industry and government data. A disadvantage of this approach is that it is likely to be expensive to establish and operate, whereas many ISACs and ISAOs are already in place and sharing some type of cyber information.

Who Collects the Data?

Another challenge is determining what type of organization—for profit, not-for-profit, or government—should perform the functions described above.

Trust and technical capability are important criteria for identifying who is best positioned to perform data collection, analysis, and dissemination. Employers will not share information if they do not believe that the data they provide will be secured and used only for approved purposes. At the same time, the collecting entity will need to do more than collect and protect cyber workforce diversity data—it will need to perform analyses and share those findings.

Considering these two factors, collection by a for-profit enterprise is the least favorable option. Companies are unlikely to share sensitive information about the diversity of their workforce with another for-profit entity. Even if they trust that entity now, there is always the risk of a merger or acquisition that could put that data in the hands of a competitor. Also, there are few laws and policies providing limits on how such data could be used and shared.

[14] "FAA Aviation Safety Information Analysis and Sharing (ASIAS)." Federal Aviation Administration. [Online]. Available: https://www.asias.faa.gov/apex/f?p=100:1::::::.

Data collection by a government entity and/or a not-for-profit struck most participants as more viable, although opinion was split. Some participants were supportive of the federal government playing an active role in gathering data, but it was unclear which government agency should be responsible for obtaining and hosting this type of information. Potential candidates include the Department of Homeland Security, the Department of Commerce, and the National Science Foundation. Other workshop participants were strongly against the notion of the federal government gathering cyber workforce diversity data. They cited four major risks:

- Wavering commitment to the effort given inevitable political transitions.
- Jurisdictional disagreements across the executive and legislative branches that fuel parochial conflicts and frustrate otherwise innocuous cybersecurity efforts.
- Fears that collected data might be used for political purposes (i.e., as a pretext for attacking certain companies) and concerns about the politicization of diversity, equity, and inclusion initiatives.[15]
- Perceptions that government-held data would be subject to Freedom of Information Act (FOIA) requests and open participating companies to unwanted public scrutiny.

By contrast, most workshop participants embraced the notion of a not-for-profit organization collecting, storing, and analyzing the data. They did so because:

- Companies would be more willing to share sensitive data with an organization that does not regulate their industry or compete in any relevant markets.
- A not-for-profit would not be subject to changes in government policy, serving as a stable and long-term partner to both industry and government.
- A non-governmental organization would not be subject to FOIA requests and could be incentivized to have strong security practices due to liability concerns and contractual arrangements that would not apply to government.
- A not-for-profit organization could operate both the ISAC/ISAO and ASIAS models of data sharing.

Weighing the various pros and cons of the different options, we believe that **reliance on a not-for-profit entity for data collection, storage, analysis, and dissemination is the best path forward**. Several types of not-for-profit organizations could do the job, including, but not limited to, industry associations, information sharing and analysis organizations/centers, think tanks, federally funded research and development organizations, and university-affiliated research centers.

Who Pays?

Finally, there is the question of funding. Cyber workforce DEI information is being gathered on an ad hoc basis by various government, industry, and not-for-profit entities (see Section 1.3). This approach is not producing the type of data that is needed to make informed policy decisions. Achieving that outcome will require data gathering and analysis on a regular basis, and the data that is gathered must be more detailed than what we've seen to date. This is unlikely to happen on its own. The type of operation that we have described in the study will likely require funding of several hundred thousand dollars per year over

[15]M. Ward, "President Biden reverses Trump's executive order banning certain diversity trainings," *Business Insider*, 21-Jan-2021. [Online]. Available: https://www.businessinsider.com/biden-reverses-trumps-executive-order-banning-diversity-trainings-2021-1. [Accessed: 07-Jan-2022].

many years. For comparison, the ASIAS program is currently budgeted at $4M per year.[16] While that is a much larger and more extensive activity than what we propose here, assuming a cost that is roughly one-tenth of the ASIAS annual budget seems reasonable, and a higher cost is not out of the question.

While at some point the benefits of the data gathering, analysis, and dissemination we propose may be large enough to prompt industry to pay for the activity, assuming this will happen is risky. It may be possible to fund an initial pilot program via civil society, but the real challenge is long-term sustainment, which is critical because it will take several years to accrue the benefits associated with this endeavor. For this reason, the most straightforward solution is for the USG to fund the long-term cyber workforce diversity program we have proposed.

First, the benefits of this initiative will support U.S. economic and security goals—a point that has been made by numerous senior government officials from both political parties.[17] Second, the USG is already heavily invested in growing the nation's cyber workforce. Expanding its focus more heavily into the DEI aspects is a natural extension of its existing mission. For example, CISA hosts a website called the National Initiative for Cybersecurity Careers and Studies (https://niccs.cisa.gov/) that provides numerous resources focused on helping organizations grow and strengthen their cyber workforces. CISA has also awarded grants to two non-governmental organizations to help develop cyber workforce programs that target under-served populations.[18] In addition, the National Science Foundation's National Center for Science and Engineering Statistics already gathers diversity data on the United States' STEM workforce and has produced a report on "Women, Minorities, and Persons with Disabilities in Science and Engineering."[19]

Finally, there is draft legislation calling for the type of data gathering that we have described in this report.[20] The National Science Foundation for the Future Act (H.R. 2225) directs the NSF Director (working in cooperation with the National Institute of Standards and Technology, the Department of Homeland Security, the Department of Defense, the Office of Personnel Management, and other federal agencies as required) to "award grants on a merit-reviewed, competitive basis to institutions of higher learning or non-profit organizations (or consortia of such institutions or organizations) to carry out research on the cyber workforce."[21]

[16]Source: Conversation with ASIAS Management Team. December 8, 2021. ASIAS funding can vary based on the number and scope of studies that are carried out each year. At this stage in its maturity, the annual cost of running of ASIAS is consistently in the millions of dollars.

[17]For example, see K. Macri, "The workforce shortage is a major cyber risk," *GovCIO Media & Research*, 29-Oct-2021. [Online]. Available: https://governmentciomedia.com/workforce-shortage-major-cyber-risk.

[18]"CISA awards $2 million to bring cybersecurity training to rural communities and diverse populations," *Cybersecurity and Infrastructure Security Agency CISA*, 20-Oct-2021. [Online]. Available: https://www.cisa.gov/news/2021/10/20/cisa-awards-2-million-bring-cybersecurity-training-rural-communities-and-diverse.

[19]K. Hamrick, "Women, Minorities, and Persons with Disabilities in Science and Engineering," *National Science Foundation*, 29-Apr-2021. [Online]. Available: https://ncses.nsf.gov/pubs/nsf21321/report.

[20]The Cyberspace Solarium Commission made a similar recommendation. See U.S. Cyberspace Solarium Commission, "Cyberspace Solarium Commission", Report, March 2020, p4. [Online]. Available: https://www.solarium.gov/report.

[21]U.S. House. 117th Congress, (2021, March 26). H.R.2225 National Science Foundation for the Future Act. [Online]. Available: https://www.congress.gov/bill/117th-congress/house-bill/2225?s=1&r=80.

This research must include an analysis of demographic representation. H.R. 2225 was passed by the House in June and sent to the Senate in July where it was referred to the Committee on Health, Education, Labor, and Pensions. No further action has been taken; its passage remains uncertain.

Recommendations

In planning a path forward, the threshold question becomes whether the benefits of collecting, analyzing, and sharing data on cyber workforce diversity outweigh the costs, which are directly proportional to ambition. A one-time survey on workforce diversity that piggybacks on preexisting data collection efforts (including longstanding industry surveys) would be relatively inexpensive. However, it would also provide limited value. In contrast, a comprehensive and detailed data gathering effort that creates a high-fidelity picture that can inform specific practices and allow organizations to steer limited resources with precision would be both highly beneficial and more costly.

A follow-on study might investigate how organizations would use cyber workforce diversity data to change hiring, training, and retention practices. The results could offer a more objective basis for weighing the tradeoffs of different approaches. However, our research and expert interviews suggest that taking the time to perform another study might be unwise. We believe that the cyber community would be best served by following a learn-by-doing approach that is biased toward action. To that end, we call on the philanthropic community, industry leaders, and appropriate government offices to consider supporting the following steps:

- Fund a year-long pilot program run by a not-for-profit organization with the appropriate technical skills and the trust of both government and industry. This organization should develop and test multiple options for data gathering, analysis, and sharing. At the end of a year, the organization will share lessons learned and proposed next steps.
- In parallel to the pilot program, put in place the mechanisms needed to sustainably fund a multi-year initiative focused on the gathering, analysis, and dissemination of DEI cyber workforce data.

We also offer three options that can be considered for the pilot program:

- **Cross the Chasm:** A focused effort with committed organizations
- **Go Big:** A broad-based effort across the entire economy
- **Split the Difference:** An industry sector-focused effort

Each of these options is described briefly below.

Cross the Chasm

There are multiple ways that one could go about addressing the challenges highlighted in this study. One option is to work with companies that are already taking steps to increase DEI. For example, the proposed data gathering and analysis organization could focus on the 31 companies that are already working with the Aspen Cybersecurity Group to "expand their aperture for cybersecurity talent" by committing to change their hiring practices in ways that will promote greater diversity.[22] The organizations that have already signed on to the ACT report are another option. Working with committed organizations offers several benefits. The participating companies will be highly motivated to work with the data collector

[22]D. Forscey and J. Purves, "16 More Industry Leaders Commit to Principles to Grow the Nation's Cybersecurity Workforce," *The Aspen Institute*, 26-Feb-2020. [Online]. Available: https://www.aspeninstitute.org/news/press-release/growing-cybersecurity-workforce/.

and are likely to provide useful feedback. In return, the participating companies would be able to:

- Determine how well their new hiring practices are working with respect to diversity.
- Assess the efficacy of retention practices, which are not currently the direct focus of the Aspen initiative but are critical for achieving the long-term goals of the effort.
- Enable the group of participating companies to learn from each other without revealing sensitive information.

The downside of this approach is that focusing on such a small set of companies limits the sample size of the pilot program. It may also produce biased results because the companies working with Aspen are already highly motivated to address DEI issues in the cyber workforce. Lessons from such a study might not scale across the nation.

Go Big

An alternative approach would focus on breadth. For example, the funded data collection and analysis organization could develop and administer a mix of surveys to the Fortune 500 and then evaluate different sharing models for its analyses. This approach would provide insights into the costs and benefits associated with a large-scale endeavor. It could build the project's credibility across a larger group of key stakeholders and potentially attract their support for a more rigorous methodology down the line. It could also support the evaluation of DEI performance by region or sector.

Of course, there are also downsides associated with this option. Companies are already being asked to answer multiple surveys, and one more survey may not go over well. As a result, participation rates might be low, and results could be biased if responses are skewed toward specific industries. Given the scope of this option, it might be difficult to communicate effectively with potential participants. Without a clear understanding of the goals of this pilot program and the potential benefits that they might receive from engaging in the effort, companies would be unlikely to engage. This could result in poor survey response rates, which in turn would make it difficult to accurately test different models for data analysis and sharing.

Split the Difference

The two ideas proposed above are not mutually exclusive: Both could be pursued simultaneously if sufficient funding were made available. There is also a middle path: The funded organization could focus on working with a single industry segment. For example, the pilot program could focus on a critical infrastructure sector where the cyber workforce shortage appears to be particularly acute. Three sectors appear promising: healthcare, finance, and the cyber industry.

The healthcare sector might be a good place to start given the dramatic increase in ransomware attacks against targets in that domain.[23] The Health ISAC could serve as a key partner due to its experience working with sector participants on different information sharing activities.[24] The financial sector faces a constant stream of cyber-attacks, and its ISAC is the most sophisticated and effective such organization in the country. The Financial Services ISAC has the technical infrastructure

[23]N. Warfield, "Why Healthcare Keeps Falling Prey to Ransomware and Other Cyberattacks," *Threatpost.com*, 02-Jul-2021. [Online]. Available: https://threatpost.com/healthcare-prey-ransomware-cyberattacks/167525/.

[24]For more information, see Health-ISAC Inc., "Crowdsourced Cyber Security | Sector Threat Intelligence | Shared Best Practices," *Health Information Sharing and Analysis Center*, Dec-2021. [Online]. Available: https://h-isac.org/.

and processes to support both information sharing and data analysis functions for its members; expanding its remit to include DEI data does not appear to be a big reach.[25]

A final option is to focus on the cyber industry itself. Workforce diversity in this sector is important because a relatively small number of companies provide most of the products and services that are used to protect our nation. If this sector's workforce were diverse, equitable, and inclusive, the benefits would ripple through the country.

[25]FS-ISAC Inc., "Safeguarding the Global Financial System by Reducing Cyber Risk," *Financial Services Information Sharing and Analysis Center*, 2022. [Online]. Available: https://www.fsisac.com/.

ACKNOWLEDGEMENTS

The author would like to acknowledge the following organizations and individuals for their contributions to and support of this report: the Hewlett Foundation, Aspen Digital, Laura Sanchez, David Forscey, Laura Bate, Marian Merritt, Sydney Jones, Cherne Lodrina, Chris Folk, Lisa Bembenick, Leslie Dunbar, Danielle Coates, Rodney Peterson, Vipin Arora, John Finamore, and all of the participants from our two workshops.

About MITRE

MITRE's mission-driven teams are dedicated to solving problems for a safer world. Through our public-private partnerships and federally funded R&D centers, we work across government and in partnership with industry to tackle challenges to the safety, stability, and well-being of our nation.

The views, opinions, and/or findings contained herein are those of the author(s) and should not be construed as an official government position, policy, or decision unless designated by other documentation.

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD®