NOVEMBER 2020

U.S. Demand for Talent at the Intersection of AI and Cybersecurity

CSET Data Brief

AUTHORS

Cindy Martinez

Micah Musser

Introduction and Summary

Cyberattacks have become more frequent, costly, and adaptive over the last decade. Despite a national demand for cybersecurity experts, the United States faces a growing deficit in skilled workers. It is estimated that the U.S. cybersecurity workforce needs to grow by 62% to meet the demands of businesses today.[1] To help close this gap, some organizations are turning to AI to augment overwhelmed cybersecurity teams.[2]

AI-enabled cybersecurity products are alluring because of their scalability and speed. When they work well, AI systems can generate security alerts based on suspicious behavior, triage potential attacks, and take some initial steps to respond to threats. This frees up analysts to focus on complex, higher priority tasks that require human intervention.

But developing AI-cyber solutions requires a workforce that can understand both cybersecurity issues and the underlying technology of AI. Although some employers may decide to build this talent by hiring distinct teams to specialize separately in AI *or* cybersecurity, there are benefits to having individuals who have some fluency in both areas. For example, AI engineers building an intrusion detection system will likely benefit if they understand not only the indicators of different attacks, but also what the attacks mean from a practitioner's perspective. Similarly, cybersecurity practitioners will likely be better positioned to evaluate, utilize, and maintain AI detection systems if they understand their underlying technology.

Because of this, the demand for employees with domain-relevant knowledge of both AI and cybersecurity provides a useful indicator for the overall level of investment and interest in AI-cyber solutions. Tracking the demand for these types of employees can help build a more complete picture of how industry is pursuing AI tools to aid in cybersecurity practice.

This paper uses data from Burning Glass Technologies, which aggregates job postings from more than 45,000 sites, to examine whether the U.S. market is signaling a meaningful shift towards jobs that call for a combination of AI and cyber skills. How many of these jobs are there, where are they located, and what do they look like?

Key takeaways from the data brief include:

- Overall demand for job postings that call for both cyber and AI skills remains low, despite considerable growth from 2010-2019.
- These job postings appear to skew more heavily toward data science roles requesting some familiarity with cybersecurity, rather than cybersecurity roles requesting some familiarity with data science.
- The finance and insurance industry is a main driver for jobs that require a combination of cyber and AI skills.
- Jobs requesting both cyber skills and AI skills are increasingly spread throughout the country, rather than concentrated in states with major AI industries.

Methodology

This research used datasets collected by Burning Glass Technologies, which scans job boards, corporate websites, and other online sites where jobs are advertised to catalog unique, active postings.[3] Important metadata is extracted and categorized into dozens of fields, such as job title, employer, and necessary skills.

To explore our research questions, we began by identifying jobs from within the Burning Glass dataset that were cyber-related and jobs that were AI-related. Because the motivation for this research emphasizes the need for workers who have both AI and cyber skills, we used Burning Glass's skill taxonomies to identify relevant jobs. These taxonomies result from searching job postings for references to hundreds of skills and then grouping them into larger skill clusters. We selected "AI jobs" by extracting any job that had at least one skill in any of the following clusters: artificial intelligence, machine learning, data science, and natural language processing. "Cyber jobs" were defined as jobs that had at least one skill in any of the following cyber-relevant skill clusters: cybersecurity, firewall software, application security, network security, internet security, anti-malware security, and information security. Simply searching by skill clusters returned a significant number of job postings that were deemed clearly unrelated after a manual review. However, we found that the vast majority of these unrelated job postings could be excluded by eliminating five career areas from both datasets.*

It is important to underscore that if a specific skill is associated with a job posting in the Burning Glass dataset, this simply means that the skill was listed somewhere in the job description; it does not mean that possessing that skill was a non-negotiable hiring criterion, nor that a candidate was required to have mastery in order to qualify. The jobs we explore in this dataset should not be treated as jobs for which employers absolutely *needed* AI-cyber "unicorns" with detailed knowledge of both domains. Many of them simply may be positions where employers indicated they would *like* applicants to possess both skills. For this reason, even if genuine AI-cyber unicorns are

---

* The excluded career tracks were: performing arts; agriculture, horticulture, and the outdoors; design, media, and writing; hospitality, food, and tourism; and healthcare, including nursing. Of these, only the last two exclusions eliminated a sizable number of jobs. We suspect that the large number of false positives identified in the healthcare industry is related to the fact that in order to comply with HIPAA regulations, healthcare workers must often have special training related to information security.

rare, our skill-based approach lets us gain insight into broader trends regarding market interest in the intersection of AI and cybersecurity.

Finally, a note on the limitations of this research: the data available to us from Burning Glass does not provide any information about whether or not job vacancies are ultimately filled. Nor, in most instances, can we access full job descriptions in order to determine exactly what type of work is being performed. Nonetheless, online job posting data is estimated to reliably capture roughly 60-70% of labor demand, and an even larger percentage for jobs requiring a Bachelor's degree or higher.[4] The use of this data is therefore a robust tool for assessing shifts in labor demand in highly professionalized domains such as AI research and the cybersecurity industry.

Findings

Over the course of the past decade, the number of job postings for both cyber skills and AI skills—which we will hereafter refer to as "overlap jobs"—has grown considerably. These overlap job postings swelled from only a few hundred in 2010 to around 34,000 in 2019, a growth of approximately 2,700%. By comparison, the categories of cyber jobs and AI jobs grew by approximately 265% and 1,300%, respectively. Nevertheless, the overall demand for employees with skills in both AI and cyber remains low when compared to the demand for employees with skills in only one discipline, as shown in Figure 1. In 2019, only about 10% of AI job postings requested competency in any cyber skills, and only about 5% of cyber job postings requested competency in any AI skills.

Figure 1. Overlap job postings have grown considerably in the last decade

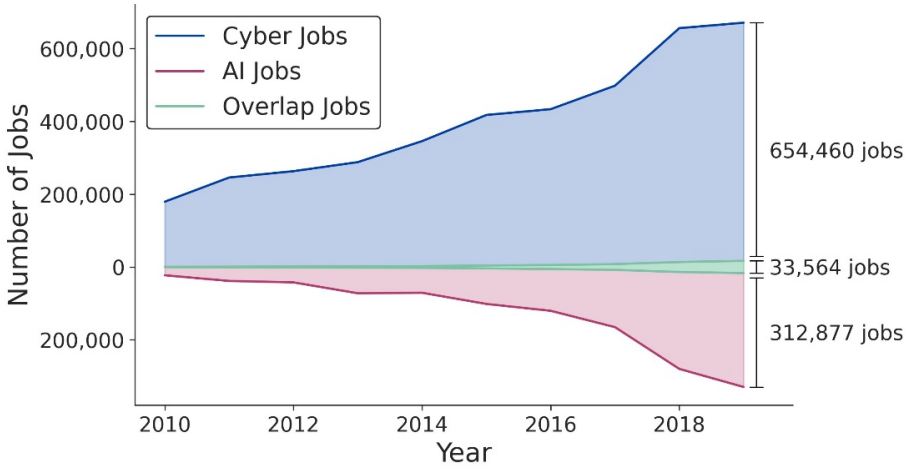

Figure 1 shows job posting growth from 2010-2019 for cyber jobs, AI jobs, and overlap jobs in the United States.

Source: Burning Glass Dataset.

To better understand the nature of overlap jobs, we began by examining their most common job titles using data from 2010-2019. Table 1 lists the 10 most common job titles for the overlap jobs in our dataset. Interestingly, when we separately examine the 10 most common AI job titles and the 10 most common cybersecurity job titles, we find that seven of the most common overlap jobs are also common AI jobs, while only three are common cybersecurity jobs. This suggests that the growth in overlap jobs is likely driven by an increasing demand for data scientists with some knowledge of

cyber, as opposed to an increase in demand for cyber professionals with some knowledge of AI.

Table 1. Job titles for overlap job postings more closely resemble AI jobs

Rank	10 Most Common Overlap Job Titles *^ Denotes a job title that appears on the list of top 10 cyber jobs* *† Denotes a job title that appears on the list of top 10 AI jobs*
1.	Software Development Engineer † ^
2.	Data Architect †
3.	Solutions Architect †
4.	Security Engineer ^
5.	Systems Engineer † ^
6.	Data Analyst †
7.	Security Architect
8.	Software Developer †
9.	Java Developer †
10.	Intelligence Analyst

Table 1 lists the most common job titles for all overlap jobs posted between 2010 and 2019, using the canon_title field available in the Burning Glass dataset, which was also used to generate lists of the 10 most common cybersecurity job titles and the 10 most common AI job titles. The 10 job titles in this table represent 24.9% of all overlap jobs.

Source: Burning Glass Dataset.

These jobs, it seems, may still be more fundamentally AI-focused than cyber-focused.[*] But how are these overlap jobs distributed? To begin answering this question, we broke down the growth in overlap jobs by industry, using the economic sector classifications of the North American Industry Classification System. Burning Glass maps the different companies in its dataset onto a

[*] A more mathematically formal comparison of the composition of these three sets of jobs can be found in the Appendix, which only serves to strengthen our claim that the overlap jobs in our dataset resemble AI jobs more closely than they resemble cybersecurity jobs.

specific NAICS code that represents the organization's primary business activity.

Figure 2 (below) shows the change in overlap job postings since 2010 for the six NAICS sectors that have demonstrated the most demand for employees with both cyber and AI skills. Professional, scientific, and technical services represent the largest source of overlap jobs. This is a broad sector that includes many jobs requiring high degrees of expertise and specialization, such as computer programming, scientific research and development, and consulting services. As a result, it is not necessarily a surprise to see this sector contain the most postings.

The industry with the second-largest demand, however, is the finance and insurance industry, which nearly quadrupled its demand for overlap jobs between 2016 and 2019. Because of their high need for data security, financial organizations are often on the forefront of new cybersecurity technologies, including AI, which has been used for years within the industry to detect potential fraud and abuse.[5]

Figure 2. The technical and financial industries were the first to begin hiring for overlap jobs

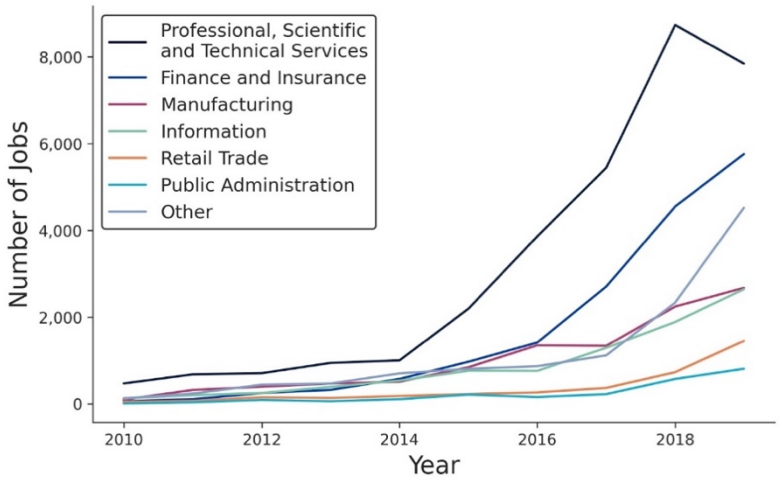

Figure 2 shows the number of overlap jobs in the U.S. from 2010-2019. The top six sectors have been identified as professional, scientific and technical services; finance and insurance; manufacturing; information; retail trade; and public administration. All remaining sectors have been grouped and labeled as "other."

Source: Burning Glass Dataset.

Figure 2 also suggests that the number of overlap jobs within the Public Administration sector has not shifted dramatically over the last decade. Yet

even if it is not directly hiring employees with AI and cyber skills, the public sector may still be a significant beneficiary of the work done by these employees. We base this conclusion on an examination of some of the top U.S. employers for overlap jobs, shown in Table 2. Many of these top employers are companies that perform significant government contracting services, which may mean that many of these employees are indirectly performing services for the public sector. The significance of the finance and insurance industry in driving the demand for overlap jobs is further corroborated by this table: five of the top fifteen employers are finance or insurance companies.

Table 2. Top 15 employers for overlap job postings

Rank	Employer	Number of Jobs
1.	CACI	3736
2.	Deloitte	2532
3.	IBM	2428
4.	Aetna Incorporated	2087
5.	U.S. Bancorp	1775
6.	Amazon	1742
7.	Booz Allen Hamilton, Inc.	1651
8.	Accenture	1220
9.	Wells Fargo	1166
10.	JP Morgan Chase	1069
11.	Verizon	1046
12.	Microsoft Corporation	1035
13.	American Express	957
14.	Raytheon	879
15.	General Electric	797

Table 2 shows the top 15 employers for overlap jobs in the United States both by the total count of overlap jobs posted 2010-2019.

Source: Burning Glass Dataset.

Finally, we looked at the way in which the geographic distribution of these jobs changed between 2015 and 2019. If AI-cyber tools are viewed by the broader cybersecurity industry as a fad or are still too immature for successful deployment, we might expect to see these jobs overwhelmingly concentrated in major AI hubs but missing from other parts of the country. To some extent, this is what we observed in 2015, as shown in Figure 3, which plots the percentage of all AI or cyber jobs that required both AI and cyber skills by state. Because a darker hue indicates more overlap between the AI and cyber industries in a given state, Figure 3(a) suggests that in 2015, significant overlap only existed in states such as California, Massachusetts, Washington, and New Mexico, most of which have well established AI hubs.[6]

Figure 3. Overlap jobs are no longer clustered solely in AI hubs

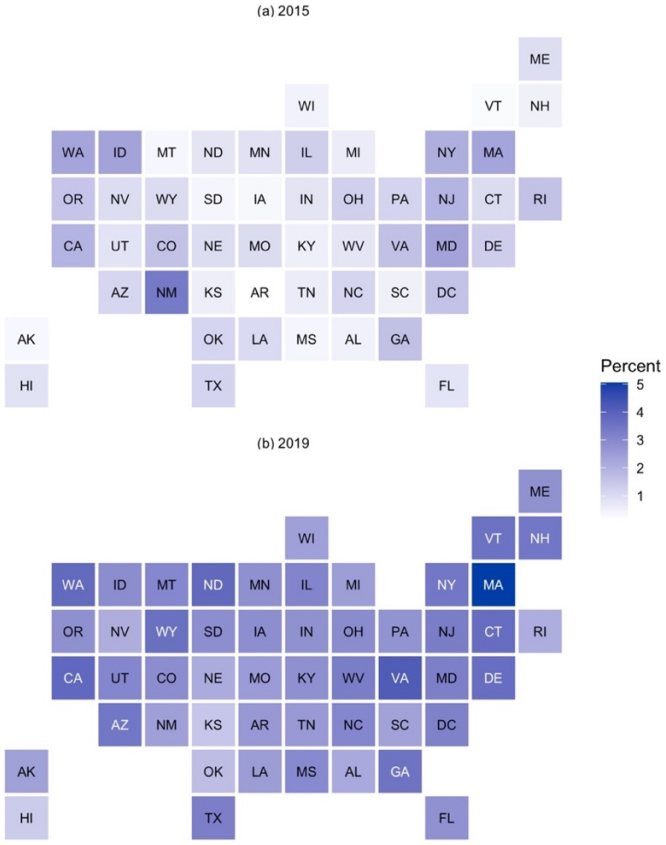

Figure 3 plots the geographic distribution of overlap jobs by calculating the percentage of jobs requiring both AI and cyber skills out of all jobs in a given state that require either AI or cyber skills. Darker hues indicate that a larger percentage of cyber or AI jobs are overlap jobs requiring both skills.

Source: Burning Glass Dataset.

Since 2015, while the degree of overlap between the AI and cyber industries of these states increased somewhat (with the exception of New Mexico), overlap jobs no longer appear to be primarily clustered in AI hubs. Figure 3(b) suggests that by 2019, the demand for overlap jobs throughout the country was comparable to or greater than the demand from major AI hubs only four years earlier. This suggests that interest in AI-cyber tools has diffused beyond the major AI hubs, which may indicate that we are observing the beginnings of an industry-wide shift towards greater utilization of AI tools.

Conclusions

The talent shortage facing the cybersecurity industry is widely noted, and media frequently invoke AI as a tool that can help address this problem through automation. But if AI will ever be a major tool in the cybersecurity arsenal, it will need to be developed, implemented, and maintained by people or teams who understand both cybersecurity and AI. This report gives reasons to be both optimistic and cautious about whether companies are taking the right steps to fully utilize the benefits of AI.

On the positive side, it is clear that there is a growing interest in finding employees who have both AI and cyber skills: since 2010, the number of job postings asking for this combination of skills has grown in both absolute and relative terms. Many of these jobs are predictably clustered in the professional and financial sectors of the economy, and it seems that demand for them is surprisingly spread out across the country, rather than being clustered solely in states with major AI industries. Based on the top job titles, these jobs seem to be principally data scientist or AI researcher jobs that are requesting cyber skills, rather than cybersecurity jobs that are requesting AI skills.

At the same time, market demand for overlap jobs remains low, which undercuts technologists' predictions that AI can substantially transform cybersecurity in the near future. While the prevalence of overlap jobs likely understates the total demand for investment in AI-cyber tools—since many companies may choose to hire AI and cyber talent separately—it seems reasonable to conclude that knowledge of AI remains relatively rare in the cybersecurity world, and vice versa. Moreover, the fact that most of the overlap jobs that do exist seem to be data science roles at their core may pose problems down the line. Product selection, integration, maintenance, and interpretation of AI systems are all areas where cyber practitioners will struggle to make informed decisions if they are unfamiliar with the basic underpinnings of machine learning systems. Yet this report suggests that relatively few companies are proactively seeking cybersecurity professionals who are also familiar with the basics of AI.

This research provides only a partial description of the growing demand for employees at the intersection of AI and cyber. Our data does not, for instance, help us determine whether labor supply is adapting to match shifts in demand. The relationship between the growth in these jobs and the public sector is also a topic that lacks clarity, and the question of how the public

sector is benefitting from AI-cyber tools either directly or indirectly is worth further study. As researchers attempt to understand the growing role of AI in cybersecurity, it will be important to continue monitoring how the labor market changes for employees with knowledge of both domains.

Acknowledgments

For their helpful discussions, comments, and input, the authors would like to thank James Dunham, Melissa Flagg, Autumn Toney, John Bansemer, and Drew Lohn. The authors are solely responsible for the views expressed in this piece and for any errors.

Document Identifier: doi: 10.51593/2020CA009

Appendix

Cosine similarity is a measure that is frequently used in natural language processing to quantify the degree of similarity between two documents. In calculating cosine similarity, each word that occurs in an overall corpus is treated as a unique dimension *i*, and each document is represented as a vector of all of the words it contains. The similarity of the two documents can then be calculated using the equation

$$\text{similarity}[A, B] = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

which yields a value that ranges between 0 (for no similarity) and 1 (which indicates that the same words have occurred an identical number of times in both documents).

This method can be adapted for comparing sets of job titles in order to determine which sets are most similar. To perform this calculation, we treat each job title as a unique dimension *i* and calculate the cosine similarity of our three sets of jobs accordingly. Using this method, we calculate the similarity between our different sets of jobs as follows:

	AI Jobs	Cyber Jobs	Overlap Jobs
AI Jobs	1.0	0.46	0.89
Cyber Jobs	0.46	1.0	0.67
Overlap Jobs	0.89	0.67	1.0

This table suggests that the degree of similarity between cyber jobs and AI jobs is relatively low. It also suggests that there is significantly more similarity between overlap jobs and AI jobs than between overlap jobs and cyber jobs (that is, similarity[overlap, AI] > similarity[overlap, cyber]).

The Burning Glass dataset also contains two fields for job titles: a clean_job_title field, which includes the actual text of the job title, and a canon_title field, which maps these job titles onto a list of "canonical" job titles. Table 1 of this data brief examined the 10 most common overlap job titles using the canon_title field. We chose to use that field because the top 10 overlap job titles represent 24.9% of all overlap jobs when using the

standardized canon_title field, but only 6.5% when using the clean_job_title field.

However, the clean_job_title field presents a more fine-grained and unaltered view of the jobs in the dataset, and using cosine similarity allows us to take all of this detail into account when calculating the similarity between different sets of jobs. When using clean_job_title instead of canon_title, our calculations for cosine similarities change in the following way:

	AI Jobs	Cyber Jobs	Overlap Jobs
AI Jobs	1.0	0.31	0.77
Cyber Jobs	0.31	1.0	0.52
Overlap Jobs	0.77	0.52	1.0

While taking a more fine-grained approach causes all similarity scores to drop, the drop is larger in the similarity between cyber and overlap jobs (0.67 → 0.52) than it is in the similarity between AI and overlap jobs (0.89 → 0.77). This bolsters our credence that the overlap jobs in our dataset genuinely are—at their core—mostly AI or data science roles, not cybersecurity jobs.

Endnotes

[1] "Strategies for Building and Growing Strong Cybersecurity Teams" ((ISC)², 2019), https://www.isc2.org/-/media/ISC2/Research/2019-Cybersecurity-Workforce-Study/ISC2-Cybersecurity-Workforce-Study-2019.ashx?la=en&hash=1827084508A24DD75C60655E243EAC59ECDD4482.

[2] "Reinventing Cybersecurity with Artificial Intelligence: The New Frontier in Digital Security" (Capgemini Research Institute, July 11, 2019), https://www.capgemini.com/wp-content/uploads/2019/07/AI-in-Cybersecurity_Report_20190711_V06.pdf.

[3] "Frequently Asked Questions," Burning Glass, https://www.burning-glass.com/about/faq/.

[4] Anthony P. Carnevale, Tamara Jayasundera, and Dmitri Repnikov, "Understanding Online Job Ads Data: A Technical Report" (Georgetown University Center on Education and the Workforce, April 2014), https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf.

[5] Adnan M. Al-Khatib, "Electronic Payment Fraud Detection Techniques," *World of Computer Science and Information Technology Journal (WCSIT)* 2, no. 4 (2012): 137–41.

[6] Melissa Flagg and Justin Olander, "AI Hubs in the United States" (Center for Security and Emerging Technology, May 2020), https://cset.georgetown.edu/research/ai-hubs-in-the-united-states/.