Trailblazing the Artificial Intelligence for Cybersecurity Discipline: A Multi-Disciplinary Research Roadmap

SAGAR SAMTANI, Department of Operations and Decision Technologies, Indiana University
MURAT KANTARCIOGLU, Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas
HSINCHUN CHEN, Department of Management Information Systems, University of Arizona

Cybersecurity has rapidly emerged as a grand societal challenge of the 21st century. Innovative solutions to proactively tackle emerging cybersecurity challenges are essential to ensuring a safe and secure society. Artificial Intelligence (AI) has rapidly emerged as a viable approach for sifting through terabytes of heterogeneous cybersecurity data to execute fundamental cybersecurity tasks, such as asset prioritization, control allocation, vulnerability management, and threat detection, with unprecedented efficiency and effectiveness. Despite its initial promise, AI and cybersecurity have been traditionally siloed disciplines that relied on disparate knowledge and methodologies. Consequently, the AI for Cybersecurity discipline is in its nascency. In this article, we aim to provide an important step to progress the AI for Cybersecurity discipline. We first provide an overview of prevailing cybersecurity data, summarize extant AI for Cybersecurity application areas, and identify key limitations in the prevailing landscape. Based on these key issues, we offer a multi-disciplinary AI for Cybersecurity roadmap that centers on major themes such as cybersecurity applications and data, advanced AI methodologies for cybersecurity, and AI-enabled decision making. To help scholars and practitioners make significant headway in tackling these grand AI for Cybersecurity issues, we summarize promising funding mechanisms from the National Science Foundation (NSF) that can support long-term, systematic research programs. We conclude this article with an introduction of the articles included in this special issue.

CCS Concepts: • **Security and privacy → Human and societal aspects of security and privacy**; • **Computing methodologies → Knowledge representation and reasoning**; **Machine learning approaches**;

Additional Key Words and Phrases: Cybersecurity, artificial intelligence, analytics, cyber threat intelligence, security operations centers, disinformation, adversarial machine learning

ACM Reference format:
Sagar Samtani, Murat Kantarcioglu, and Hsinchun Chen. 2020. Trailblazing the Artificial Intelligence for Cybersecurity Discipline: A Multi-Disciplinary Research Roadmap. *ACM Trans. Manage. Inf. Syst.* 11, 4, Article 17 (December 2020), 19 pages.
https://doi.org/10.1145/3430360

This material is based upon work supported by the National Science Foundation under Grant Numbers OAC-1917117 (CICI), CNS-1936370 (SaTC CORE), CNS-1850362 (CRII SaTC), and DGE-2038483 (SaTC-EDU).

Authors' addresses: S. Samtani (corresponding author), Department of Operations and Decision Technologies, Indiana University, 1275 E. 10th St., Bloomington, Indiana 47405; email: ssamtani@iu.edu; M. Kantarcioglu, Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080; email: muratk@utdallas.edu; H. Chen, Department of Management Information Systems, University of Arizona, 1130 E. Helen St., McClelland Hall 430, Tucson, AZ 85721; email: hsinchun@arizona.edu.

2158-656X/2020/12-ART17 $15.00
https://doi.org/10.1145/3430360

1 INTRODUCTION

The regularity of devastating cyber-attacks has made cybersecurity a grand societal challenge. Innovative solutions to tackle ever-evolving threats in cyberspace are essential for robust cybersecurity postures. Artificial Intelligence (AI) holds significant promise in sifting through large volumes of heterogeneous cybersecurity data with unprecedented efficiency and effectiveness [25]. These benefits can lead to significant enhancements in prevailing cybersecurity tasks, including asset identification, vulnerability management, emerging threats detection, and control deployment. Despite initial successes, there remains a significant dearth of work examining how AI can be deployed for cybersecurity. The lack of growth in this area is likely attributable to the diversity, complexity, and rapidly evolving nature of AI and cybersecurity. AI draws from math, biology, and other disciplines, while cybersecurity relies on knowledge of protocols, risks, and more. Moreover, cybersecurity data has unique properties based on its underlying generating processes (humans and machines), rapidly evolving nature, and sheer volume. As a result, representing and processing such data in a manner that maximizes performance and practical utility is a nontrivial technical and nontechnical task.

Despite these significant challenges, globally recognized entities such as the National Academies of Sciences (NAS) and the National Science Foundation (NSF) have underscored the critical need for AI for Cybersecurity research [20, 26]. Significant advances in this nascent discipline can usher in a new generation of cyber resilience against an ever-evolving threat landscape. In light of these significant needs, we aim to provide a systematic overview of fundamental cybersecurity data, prevailing AI for Cybersecurity application areas, and key limitations in the extant landscape. Taking these together, we develop and propose a multi-disciplinary roadmap to rapidly advance the AI for Cybersecurity discipline. We frame the contributions of this article as follows:

(1) We summarize prevailing cybersecurity data sources commonly used in extant AI for Cybersecurity research. The review broadly categorizes data sources available within and outside of any particular organization.
(2) We summarize four major themes of AI for Cybersecurity within the larger scholarly and practitioner communities: cyber threat intelligence (CTI), security operation centers (SOCs), disinformation and computational propaganda, and adversarial machine learning (ML). For each theme, we provide an overview of the data it primarily relies on, as well as academic and industry pioneers.
(3) We provide an end-to-end and integrated roadmap that scholars and practitioners can follow when aiming to conduct the next generation of AI for Cybersecurity research. This roadmap intentionally emphasizes a multi-disciplinary perspective and provides concrete examples of promising future directions for research.
(4) We provide a summary of prevailing grant funding opportunities, conference venues, and journal outlets that AI for Cybersecurity scholars and practitioners can consider when aiming to grow the discipline. Presenting a consolidated list of resources in this fashion can help the AI for Cybersecurity community rapidly build sustainable, visible, and highly-impactful research and education.

This remainder of this article is organized as follows. First, we provide a comprehensive overview of prevailing cybersecurity data sources. Second, we summarize a past and present view of AI for Cybersecurity initiatives that rely on these data. Third, we highlight some of the key gaps within the existing academic and industry research landscapes and present our integrated AI for Cybersecurity research roadmap. Fourth, we summarize prevailing NSF funding opportunities that

can support AI for Cybersecurity research. The final section presents the articles in this special issue and concludes this work.

2 PREVAILING CYBERSECURITY DATA SOURCES

Executing effective AI for Cybersecurity is contingent upon analyzing rich data sources. Two broad categories of cybersecurity data exist: internal and external. Internal data pertain to resources available within an organization. External cybersecurity data refer to content accessible to the broader public (i.e., outside of an organization). We provide a summary of prevailing data sources in each category in Table 1. For each data source, we provide a brief description, example platforms, and sample metadata.

Internal data hold tremendous value for developing cybersecurity as they are close to an organization's critical assets. Most critical assets (e.g., web servers, databases, embedded systems, routers, etc.) are based on workstations and/or virtual machine (VM) images. These devices will often possess significant data that can help a hacker and/or a systems administrator identify their contents. Sample metadata include operating system (e.g., Windows, Linux, Unix, etc.), version, file system structure (e.g., size, quantity, file names), and others. Each networked workstation and VM generates data that can help a systems administrator fingerprint its characteristics. Depending on how the workstation and VM are configured, they may also generate alerts (e.g., events, timestamps, etc.) about selected activities occurring on the device. Inter-connected devices on a network can often have significant netflow that contains data such as source, destination, bytes, headers, and others. Vulnerability assessments conducted by tools such as Nessus, Qualys, OpenVAS, Burp Suite, and others can help systems administrators and security analysts detect the flaws of technologies deployed across their network. Biometric data provide sensor readings of human (e.g., employee actions) within a network. Finally, enterprise networks may also contain intranets equipped with social media sites, collaboration tools (e.g., Slack, Teams, Confluence, etc.), and internal reports.

While internal network data can provide low lead time when aiming to provide knowledge about existing threats and past attacks, external data sources can help facilitate knowledge about events occurring in broader cyberspace. Social coding repositories (e.g., GitHub, SourceForge, etc.) and Internet-of-Things Search Engines (IoTSEs) can help an organization understand the scope of how publicly accessible their code bases are (e.g., containing technology names, usernames, passwords, etc.) and openly available devices are, respectively. Dark Web platforms can provide an understanding of the online hacker community and their relevant tools, techniques, and processes. For example, hacker forums provide millions of freely accessible exploits and facilitate discussions to allow hackers to execute cyber-attacks. DarkNet Marketplaces (DNMs) and carding shops provide mechanisms to sell illicit goods (e.g., stolen credit cards) to reap financial benefit. Internet-Relay-Chat (IRC) channels are often used by hacking groups such as Anonymous to discuss the targets of their breaches. Selected plain-text Dark Web contents found on hacker forums, DNMs, carding shops, and IRC channels may sometimes be available in Paste Sites (e.g., PasteBin) too. Commercial threat feeds provided by prevailing CTI companies are designed to help industries be aware of prevailing threat trends. Finally, news source and conventional social media (e.g., Twitter, Facebook, YouTube, etc.) can provide knowledge about the key security issues afflicting society as a whole.

3 A PAST AND PRESENT VIEW OF AI FOR CYBERSECURITY APPLICATION AREAS

When taking internal and external data sources together, an organization can have significant metadata (e.g., timestamps, author names, etc.) and data (e.g., rich text content, risk score, etc.) to execute fundamental cybersecurity tasks from an AI-based perspective. Recognizing the significant promise and potential of such a cybersecurity big data gold-mine, numerous scholars and

Table 1. Summary of Prevailing Cybersecurity Data Sources

Type	Data Source	Description	Example Platforms or Tools	Sample Metadata and Data
Internal	Workstations and/or virtual machine images	Machines that enable and facilitate computational activities	Docker, containers, VMware	Operating system, applications, file systems
	Data storage	Devices that store data from users and networks	File store, disk drives, file directories	File size, directory name, file name, directory size
	Networking devices	Devices that help route and facilitate network traffic	Routers, switches, gateways, SDN software	ARP, routing tables
	Network-based fingerprint data	Data that pertains to the content that a device generates	Nmap, Zmap	TCP, packer header, UDP
	Netflow data	Flow of data across networked devices	–	Source, destination, bytes
	Alert and event logs	Succinct reports and short alerts about selected network and/or device events	Security information and event management systems (SIEM)	Date, alert name, IP address, action
	Vulnerability assessment	Reports generated from prevailing vulnerability scanning tools	BurpSuite, Nessus, Qualys, OpenVAS	Name, severity, risk
	Biometric data	Data generated from sensors that monitor human behaviors	Mouse movements, eye movements, pulse	x-y-z axis accelerometer readings
	Intranets and social media, and reports	Tools to facilitate collaborations across teams	SharePoint, Confluence, Teams, Slack	Usernames, plain text, multi-media data
External	Social coding repositories	Sites that enable the sharing of code in repositories	GitHub, SourceForge	Commits, authors, code, forks
	Internet-of-Things (IoT) search engines	Search engines that search and index publicly accessible IoT devices	Shodan, Censys, Fofsa, BinaryEdge,	IP, banner data, images, latitude, longitude
	Hacker forums	Online discussion boards that allow hackers to discuss malicious attacks	Antichat, Ciphers, WildersSecurity, go4expert	Date, author, threads, source code
	DarkNet marketplaces	Markets that facilitate the sale of illicit goods	Hansa, DreamMarket	Product name, author name, price
	Internet-Relay-Chat	Plain-text instant messaging chatrooms often used by hacktivist groups	Anonops	Date, plain-text
	Carding shops	Sites that sell stolen credit cards	JStash, Recator	Card type, zip code,
	Paste sites	Sites that allow anonymous posting of plain-text content	PasteBin	Raw paste, author, date, size
	Commercial Threat Feeds	Feeds of threat intelligence data curated by industry	AlienVault OTX	IPs, hashes, source, destination
	Malware Repositories	Sites that aggregate malware and reports	EMBER, VirusTotal	Hash, binary, date, malware reports
	News sources	Public media sources that share news about events	CNN, BBC, Fox, ABC	Headlines, text bodies, images
	Conventional social media	Sites that facilitate social networking	Twitter, Facebook, YouTube	Usernames, plain text, multi-media data

Table 2. Summary of Prevailing AI for Cybersecurity Application Areas

Application Area	Selected Common Tasks	Selected Accessible Datasets	Selected Tools	Academic Pioneers	Industry Pioneers
CTI	Malware analysis	VirusTotal	Cuckoo	UTD CS	FireEye
	Phishing detection	PhishTank	PhishMonger	UArizona AI Lab	KnowBe4
	Dark Web Analysis	AZSecure HAP	ISILinux		CYR3CON
Disinformation and Computational Propaganda	Bot detection	Bot Repo, Twitter Bot-Cyborg	Hoaxy, Botometer	Computational Propaganda Project	Paragon Science
	Disinformation identification	Credibility Coalition, GOP Twitter	Exifdata, exiftool, factcheck	CMU Center for Informed Democracy	CarleyTech, Rand, FireEye
SOC	Log file analysis	Boss of the SOC	Kiwi, Splunk	UC Irvine CS, U. Michigan CS, R. Marty	Splunk
	Vulnerability assessment	NVD, Metasploit	Nessus, ZMap		Tenable
	Intrusion detection	CIC-IDS 2017	Zeek		Palo Alto
Adversarial ML	Malware evasion, ML poisoning	EMBER, NIPS Adv. learning	EvadeML, SecML	Ian Goodfellow, Nicholas Carlini	Elastic, Google Brain, Microsoft

*Note: CS = Computer Science; CMU = Carnegie Mellon University; GOP = Grand Old Party; HAP = Hacker Assets Portal; NVD = National Vulnerability Database; NIPS = Neural Information Processing Systems; ISILinux = Intelligence and Security Informatics Linux; UArizona = University of Arizona; UC = University of California; UTD = University of Texas, Dallas.

practitioners have started leveraging AI for four data-rich and ever-evolving cybersecurity applications: (1) CTI, (2) disinformation and computational propaganda, (3) SOCs, and (4) adversarial ML. Each application area relies on significant quantities of varying data presented in Table 1. In Table 2 we provide a detailed summary of each AI for Cybersecurity application area, selected common tasks, datasets, tools, academic pioneers, and industry pioneers.

While each application area of AI for Cybersecurity is presented separately in Table 2, they all have some overlap. For example, SOC analysts often contend with adversarial ML issues while simultaneously ingesting CTI. In the following sub-sections, we further describe each application area, with an emphasis on describing the application area, key tasks, data source, and prevailing tools and pioneers.

3.1 Cyber Threat Intelligence (CTI)

CTI is concerned with identifying emerging threats and key threat actors to enable effective cybersecurity decision making [33]. CTI has traditionally been tightly linked with SOCs and has therefore relied on internal data sources such as netflow data, virtual machine images, vulnerability assessment, fingerprint data, and others. However, recent years have seen a significant expansion to external data sources such as hacker forums, DarkNet Marketplaces, carding shops, IRC channels, public news sources, and others. Conventional analytics tasks include threat modeling, malware analysis, IP reputation services, summary statistics, cyber-forensics, and threat hunting [6].

Despite the prevalence of these approaches, results are often high-level overviews of the data contents, rather than fine-grained insights into the patterns pervading each dataset. As a result, CTI scholars and industry organizations are increasingly turning to AI for critical tasks such as emerging threat detection and mitigation, key hacker identification and attribution, and others. Academic CTI pioneers include UT Dallas for malware analysis and University of Arizona's AI Lab for Dark Web analytics [3, 8, 12, 17, 35–37] and phishing analysis [2]. Industry pioneers include

FireEye, KnowBe4, CYR3CON, and others [34]. Numerous tools to support various aspects of CTI have been generated from these pioneers, including ISILinux for fundamental CTI analytics for the Intelligence and Security Informatics (ISI) community, Phishmonger to support advanced phishing analytics, AZSecure for Dark Web analytics, and Cuckoo and VirusTotal for dynamic and static malware analysis, respectively.

3.2 Disinformation and Computational Propaganda

Disinformation and computational propaganda research aim to examine how fake or false information propagates through international cyberspace and major geo-political regions (e.g., Russia, China, US, Middle East, etc.). Computational propaganda examines how the use of algorithms, automation, and big data approaches can influence and shape public opinions related to social and political issues. If left unchecked, disinformation and computational propaganda pose a significant threat to quickly destabilize governments and sway public perceptions and actions on societally relevant events (e.g., elections). As a result of their potentially far-reaching implications, disinformation and computational propaganda topics have gained significant traction within academic and practitioner circles. Common data sources that facilitate research focus on conventional social media platforms such as Facebook, Twitter, Reddit, Weibo, and YouTube.

While the initial methodologies employed were qualitative and based in rumor theory and criminological perspectives, recent research has started incorporating AI-based techniques such as text mining, network science, image recognition, and neural networks that automatically sift through large quantities of social media data (e.g., text, images, videos, etc.) to pinpoint computational propaganda, disinformation content, campaigns, astroturfing, bots, message amplification, fake news, and other related phenomena. Prevailing academic entities leading innovations in this space include Carnegie Mellon University's (CMU's) Center for Informed Democracy, Indiana University's (IU's) Network Science Institute, and the University of Oxford's Computational Propaganda Research Project. Industry pioneers include FireEye and Symantec Threat Intelligence. These focused efforts have enabled various resources pertaining to disinformation and computational propaganda research to emerge, including the Botometer and Hoaxy tools and datasets such as bot repo and Twitter Bot dataset.

3.3 Security Operations Centers (SOCs)

SOCs are often the heart of many organizations' cybersecurity efforts. SOCs have conventionally relied on human analysts to help to ensure the confidentiality, integrity, and availability (CIA) of selected enterprise information technology (IT) and information systems (IS) are in accordance with the larger organizational goals and industry-specific policies (e.g., GDPR, CCPA, Sarbanes-Oxley, FISMA, etc.). Key IT and IS that are often monitored include websites, applications, databases, networking technologies (e.g., switches, routers, software defined networks, etc.), mobile devices, and Internet of Things (IoT) devices. SOC analysts often rely on systems such as Network Intrusion Detection Systems (NIDS), Network Intrusion Protection Systems (NIPS), Security Information and Event Management Systems (SIEMs), Security Operations and Response (SOAR) platforms, antivirus, firewalls, and unified threat management (UTM). Such systems capture various data from sources including network traffic flow analysis, vulnerability assessment, anomaly detection, blacklisting, detecting phishing attacks and campaigns, identifying correlations, threat modeling, alert management, and others.

Despite the prevalence of these practices, extant SOC operations have been known to be prone to significant information overload, false positives, and false negatives. These issues can cause alert fatigue in security analysts and result in significant burnout and mental health issues. In recognition of these issues, AI has started to emerge as a viable approach to cut through the noise,

curtail alert fatigue, and deliver filtered results. Prevailing academic pioneers focusing on SOC research include the University of California, Irvine (UCI); the University of California, Santa Barbara (UCSB); and the University of Michigan (UMich). These entities have generated significant tools and datasets such as Zeek and the CIC-IDS 2017. Each resource is designed to support various SOC-related inquiries. Industry pioneers include Splunk, Tenable, Palo Alto, and Rafael Marty [22, 24, 29, 40, 42]. In addition to releasing tools (e.g., Metasploit) and datasets (e.g., NVD), industry pioneers have launched competitions to help encourage community involvement and enhance interest in the topic. An example of such an event is Splunk's Boss of the SOC (BOTS) event that draws hundreds of participants annually to engage in hands-on, self-paced blue-team exercises to hunt and defeat threats within networks.

3.4 Adversarial Machine Learning (ML)

The rapidly increasing popularity of adversarial ML is largely attributable to algorithms and technologies such as generative adversarial networks (GANs) [15]. Adversarial ML is an emerging class of machine learning algorithms that aim to fool algorithms by generating and/or supplying deceptive input that is strikingly similar to real data. Adversarial machine learning can be used for both offensive and defensive purposes. Prevailing offensive adversarial ML tasks include generating deep fakes (e.g., synthetic text, images, and videos), poisoning, AI-system attacks, and others. These offensive measures pose a new variation of cyber-attacks beyond conventional methods (e.g., malware, distributed denial of service, etc.) and have resulted in significant issues in AI security, trust, privacy, and dependability. Defensive adversarial ML tasks include threat modeling, attack simulation, countermeasure designs, noise detection, and evasion [7]. Prevailing methodologies within this area of AI for Cybersecurity include GANs, reinforcement learning, and actor critic networks. These approaches can learn from limited training data, closely mimic a human's learning process, and rapidly evolve to dynamic environments.

Adversarial ML is a relatively younger application area in AI for Cybersecurity when compared to CTI, disinformation, and SOC. Nevertheless, significant investments into developing adversarial ML capabilities have been made in recent years from academia and industry alike. Examples of academic pioneers include Ian Goodfellow (Stanford University), Nicholas Carlini (University of California, Berkeley), and Hyrum Anderson (University of Washington). Industry leaders include Microsoft, Google Brain, and Elastic. Prevailing tools and datasets to support adversarial ML research include SecML, EvadeML, and EMBER. Like the SOC community, adversarial ML academic and industry pioneers have also launched various competitions such as the annual malware evasion challenge.

4 A MULTI-DISCIPLINARY ROADMAP FOR AI FOR CYBERSECURITY: DATA, ANALYTICS, AND AI-ENABLED DECISION MAKING

Despite tremendous progress in each of the aforementioned areas of AI for Cybersecurity research, there remain four categories of significant issues. When taken together, these drawbacks can significantly limit the scope, scale, and impact of relevant AI for Cybersecurity research.

- First, industry and academia often operate in siloes. For example, academics often develop highly specialized solutions on older datasets that may not be representative of what is seen in practice. While attaining excellent performance in lab environments, they often suffer in production environments. Conversely, industry professionals have a tremendous amount of existing data but apply standard algorithms. These approaches often do not take into account of the unique characteristics of cybersecurity data and can lack rigorous evaluation processes commonly seen in academia.

- Second, and relatedly, there is a lack of publicly accessible, relevant, and realistic datasets. Ultimately, this results in academics performing analytics on single datasets (e.g., malware binaries, netflow data), rather than executing on multiple datasets simultaneously (much less realistic yet very useful for practical applications).
- Third, model sharing and interpretability are key concerns. Models developed in academia or industry rarely see airtime in the other's grounds. Moreover, models are often developed without consideration to how end-users would operate them (e.g., lack transparency). However, this type of detail and sharing is critically important to quickly develop relevant and timely models that can be deployed across environments.
- Finally, many individuals may lack the resources to get started or execute AI for Cybersecurity research. Therefore, providing resources to interested individuals to facilitate innovation in this space is critically needed. Doing so can rapidly accelerate the rate of AI for Cybersecurity innovation and development.

The limitations summarized above necessitate novel approaches for representing cybersecurity data, methodologies to support cross-cutting, inter-disciplinary, and high-impact AI for Cybersecurity research. As a result, a clear, crisp, and end-to-end roadmap on promising future directions is critically needed to ensure the long-term viability of the discipline. To this end, we identify three major themes for future AI for Cybersecurity research: cybersecurity applications and data, advanced AI methods for cybersecurity, and AI-enabled decision-making. Each area can be significantly enhanced by incorporating multi-disciplinary perspectives, including those from socio-technical, organizational, regulatory, cultural, cognition, and psychology disciplines. Presenting a roadmap in this fashion has significantly enhanced the focus of research around major topical areas [1, 4, 9]. Moreover, taking a multi-disciplinary approach is critical for ensuring that tackling societal issues is done in a holistic, all-encompassing manner [7]. Figure 1 illustrates each area's major components as well as the synergistic relationship between each area. We describe each area in further detail in the following subsections.

4.1 Cybersecurity Applications and Data

Progressing the AI for Cybersecurity discipline requires a strong foundation of application areas and data sources. We group emerging cybersecurity applications and data into three major groups: (1) emerging application areas, (2) emerging data sources, and (3) refined data representations. Each is described in the following sub-sections.

4.1.1 Emerging Application Areas. Prevailing application areas highlighted in previous sections have traditionally centered around enterprise IT environments. However, society is increasingly relying on technologies outside of the conventional networked workstation perspective. Examples include scientific cyberinfrastructures (e.g., science gateways), remote technologies and collaboration tools, sensor-based environments for smart homes and health, and industrial environments that deploy Supervisory Control and Data Acquisition (SCADA) Industrial Control System (ICS) or Cyber-Physical System (CPS) technologies. Each of these environments has unique combinations of data, tasks, and requirements that necessitate novel AI for Cybersecurity approaches. Examples include analyzing physical security constraints (e.g., biometric analysis) for secured facilities containing confidential information, social engineering-based susceptibilities (e.g., insider threats), and linking to industry- or government-specific risk frameworks (e.g., Open Science Risk Profile, MITRE ATT&CK, NIST, etc.) and threat models (e.g., diamond models, cyber kill chains, etc.). Each area has its own key considerations and requirements [13, 14, 18]. Future AI for Cybersecurity research can also execute analytics at varying levels of granularity. Examples include developing approaches at the macro-level (e.g., industry-wide), meso-level (e.g.,

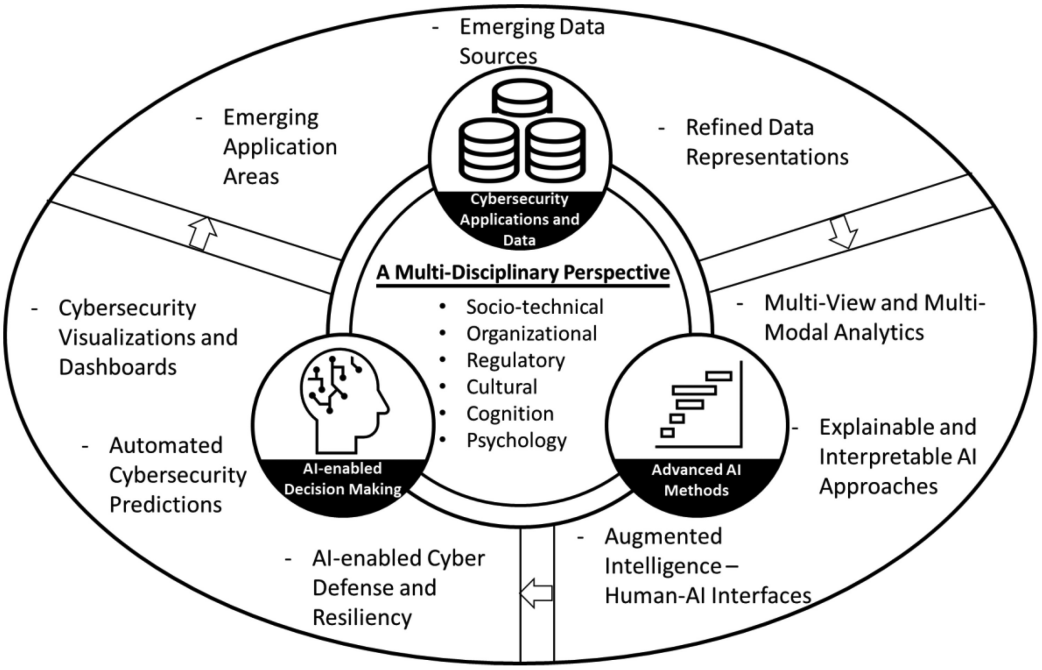

Fig. 1. A multi-disciplinary AI for Cybersecurity roadmap: cybersecurity applications and data, advanced AI methods, and AI-enabled decision making.

organization-specific), and micro-level (e.g., personalized security recommendations). Each level of granularity can help to unlock the maximum value of AI across multiple stakeholder groups.

4.1.2 Emerging Data Sources. As indicated in Table 1, extant cybersecurity data sources can be grouped into internal and external categories. While these categories encompass the breadth of data available to support various cybersecurity research inquiries and applications, data sources such as netflow, networked devices, fingerprinting, Dark Web, and social media have been leveraged quite extensively. However, data sources such as user-generated contents within an organization, biometric data, IoTSEs, and public coding repositories have received far less attention. These data sources can be leveraged in critical AI for Cybersecurity tasks. For example, examining user-generated contents (e.g., quarterly calls, internal social media, public websites, etc.) can help drive the next generation of insider threat detection and social engineering attack identification. Similarly, examining the code employees post in publicly accessible social coding repositories can help identify potential vulnerabilities (e.g., insecure coding practices, posting of private keys, etc.) that hackers can leverage to gain a foothold into an organization. Checking the consistency, correctness, and completeness of multiple IoTSEs (e.g., Shodan, Censys, Binary Edge, etc.) simultaneously can help organizations (including nonenterprise IT) effectively map out their potential attack surface and drive targeted vulnerability assessments [38]. Biometric data such as keystroke dynamics, touchscreen dynamics, eye movements, pulse rates, and others can help facilitate advanced physical cybersecurity tasks [23, 46, 47].

4.1.3 Refined Data Representations. Data representation is critical for strong algorithmic performance. To date, the prevailing approach for representing cybersecurity is a flattened feature vector. Despite its popularity, this approach omits crucial relationships apparent within the data

it represents (e.g., sequences). As such, this representation can lead to significantly diminished results when deployed in production environments. To mitigate this issue, future AI for Cybersecurity scholars can carefully consider how the cybersecurity data exists within the environment they are interested in and carefully select an appropriate option that most closely represents the phenomena of interest. For example, applications within a virtual machine can be represented as a graph (capturing their dependencies) and file systems can be represented as a tree (given their hierarchical nature). Other candidate representations include grids, sequences, and non-Euclidean (e.g., tensors, cubes). Selection of an appropriate representation can also be guided based on relevant social-behavioral economic (SBE) theories, organizational requirements, and key data characteristics [39].

4.2 Advanced AI Methods for Cybersecurity

The above-listed application areas, data sources, and data representations will require advanced methodologies to fully uncover their potential. Among various options, three major emerging methodologies can offer significant value to developing practically relevant and usable AI for Cybersecurity: multi-data source analytics, Explainable AI (XAI), and augmented intelligence (human-AI) interfaces. Each is described in further detail below.

4.2.1 Multi-View and Multi-Modal Analytics. A key drawback of the extant AI for Cybersecurity research and practice landscape is leveraging single datasets in a siloed manner. This often is a result of lack of access to multiple datasets (commonly seen within academia) and/or a thorough understanding of the relationships across multiple datasets. Not processing multiple datasets simultaneously can result in an incomplete appraisal of an environment. To address this issue, future AI for Cybersecurity research can aim to leverage the characteristics of multiple data sources in a more holistic fashion. Selected promising approaches include deep-learning-based entity matching, short text matching algorithms (e.g., deep structured semantic models), multi-view approaches (e.g., multi-source), and multi-task learning strategies [10]. Leveraging knowledge across multiple datasets via transfer learning and/or federated learning can also be employed to increase task performance [43–45]. Each model can be significantly extended to account for key domain considerations (e.g., timeliness, interpretability, etc.) and increase the model's capacity to learn. Successfully fusing multiple data sources can result in novel derived attributes, enhanced risk management scores (e.g., vulnerability assessment scores), and ultimately a holistic view of an organization's cybersecurity posture.

4.2.2 Explainable and Interpretable AI Approaches. Cybersecurity is a domain where it is critical to know how and why an algorithm reached its output decision. Unfortunately, prevailing AI-based algorithms rely on deep learning. While providing unprecedented performance in high-impact cybersecurity applications such as Dark Web analytics, vulnerability assessment, and others, they are notorious for their "black-box" nature. Lack of model explainability and interpretability can adversely affect model performance and reduce algorithm trustworthiness, security, privacy, and adoption. These drawbacks significantly hinder key cybersecurity stakeholders from effectively leveraging AI-based technologies for critical tasks. To minimize these drawbacks, future AI for Cybersecurity research can explore how interpretable and explainable AI can enhance algorithm performance as well as open their black-box nature. Two major categories of explainable and interpretable AI approaches exist: post hoc and intrinsic [11, 31]. Both categories can operate at the global or local level. Post hoc global approaches interpret major model components after they have been trained. Similarly, post hoc local approaches examine individual model processes and components (e.g., neuron activations) after the model has been trained. Intrinsic global approaches incorporate major model components directly into the architecture.

Finally, intrinsic local approaches incorporate components into a model (e.g., attention mechanisms) to identify which data features within a data input help a model reach its end output. Future research can explore how each category of the aforementioned approaches can be leveraged to extract additional insight from AI-based methodologies to support selected cybersecurity tasks.

4.2.3 Augmented Intelligence – Human-AI Interfaces. Many cybersecurity professionals would argue that AI-based algorithms and systems should not solely make cybersecurity decisions. Rather, AI-based approaches should be closely linked with human action (e.g., a security analyst is an active member in the analytics process) to help facilitate enhanced decision-making processes. Also referred to as augmented intelligence or human-AI interfaces, these approaches can lead to significant performance gains over using an algorithm or human individually. Three broad approaches of human-AI interfaces exist: substitution (AI replaces humans), augmentation (AI and human synergistically augment each other), and assemblage (humans and AI are dynamically convened to cooperate and function as a single, integrated unit) [19, 21, 32]. The breadth, scope, and depth of how humans and AI can interface for critical and fundamental cybersecurity tasks is understudied, yet critically needed. Such research would inherently need to take a multidisciplinary approach, particularly emphasizing perspectives from cognitive science, psychology, human computer interaction, and other areas.

4.3 AI-Enabled Cybersecurity Decision-Making

The results attained from advanced AI methodologies can enable unprecedented cybersecurity decision making. Three key areas that can glean benefit include cybersecurity visualizations, cybersecurity predictive analytics, and AI-enabled cyber defense. We summarize each in turn in the following sub-sections.

4.3.1 Cybersecurity Visualizations and Dashboards. Cybersecurity is inherently a data-rich, information-poor domain. While prevailing AI-based methods can help sift through tremendous amounts of noise, scholars and practitioners who work with new and/or ever-evolving cybersecurity data require a mechanism to quickly understand a dataset's key characteristics and key patterns at multiple levels of granularity. Cybersecurity visualizations are an indispensable tool to facilitate these tasks. However, most visualizations for cybersecurity are designed by security professionals who may not know about visualization theory or are created by visualization experts who lack knowledge about the nuances of cybersecurity. Therefore, there is a critical need for AI for Cybersecurity academics and professionals to critically examine how visualizations can be carefully designed to be incorporated into their workflow. Doing so can provide excellent reporting mechanisms, guide the selection of predictive algorithms, and support strategic cybersecurity functions (e.g., investments, quarterly reports, etc.). Visualizations can be conducted at the macro (global), meso (local), and micro (individual) levels [5, 41]. Visualizations can be temporal, tree, network, charts, tables, and geo-spatial, depending on the data type and key cybersecurity requirements. Visualizations can be designed to incorporate key concepts of overview, zoom, filter, history, and details on demand as well as various layouts and color schemes. Incorporating such visualizations into systems (e.g., human-AI hybrids) can help facilitate unprecedented cybersecurity decision making.

4.3.2 Automated Cybersecurity Predictions. While cybersecurity visualizations can assist in understanding the key aspects of a particular dataset, predictive analytics can help facilitate automated decision making in addressing common cybersecurity tasks such as data triage, spam filter, vulnerability classification, and mission mapping. Despite the prevalence and clarity of these tasks, many prevailing methods often lack automation or are rife with false positives. Employing

advanced AI-based predictive analytics processes can help address these selected issues. Promising methods for predictive analytics include deep Bayesian forecasting, burst detection, deep generative modeling with temporal constraints, temporal-based graph neural networks, and others. Each approach can be enhanced by including industry-specific guidelines, data, tasks, and requirements. Selected stakeholders who can benefit from enhanced predictions include SOC analysts and CTI professionals, specifically those operating at the tactical and operational levels within their organization.

4.3.3 AI-Enabled Cyber Defense and Resiliency. Organizations can leverage the knowledge gleaned from their analytics procedures to automatically deploy appropriate security controls. Examples include automated network segmentation and reorganization, automated threat modeling, heal networks following cyber-breaches, and automated patching and remediation and mitigation. Automating each task in an intelligent fashion can significantly assist SOC analysts and operators. Selected emerging AI methodologies supporting these tasks include enhanced AI agents, reinforcement learning, actor critic networks, selected defensive adversarial learning methods, and Bayesian networks. Future research can explore how each methodology can produce appropriate cyber-defenses as required across multiple vertical domains (e.g., enterprise IT, scientific cyberinfrastructure, sensor-based environments, etc.).

5 MECHANISMS TO FACILITATE ADVANCES IN THE AI FOR CYBERSECURITY DISCIPLINE

A key element to executing AI for Cybersecurity research is attaining sufficient resources (e.g., human capital, computation infrastructure, etc.) to effectively tackle emerging topics and significantly advance knowledge. In light of the significance of AI for Cybersecurity, numerous funding agencies at the federal, national, and local levels have released highly-visible grant solicitations. These solicitations often encourage researchers to develop interdisciplinary teams across technical fields such as information systems, computer science, electrical and computer engineering, information science, and social sciences as well as social-science-based disciplines such as communications, criminology, psychology, and cognitive sciences. Attaining grant funding can help these teams recruit high-caliber research scientists, foster industry and government collaborations, generate a strong reputation, and create systemic long-term research programs centered around critical AI for Cybersecurity research topics (as opposed to ad hoc teams) [16, 27, 28, 30]. Taken together, these benefits can facilitate AI for Cybersecurity innovations and discoveries at an unprecedented rate.

We provide a summary of selected NSF solicitations that have relevance to AI for Cybersecurity in Table 3. For space considerations, we only list opportunities available through the NSF. Funding from the NSF is often referred to as the "gold-standard" of funding, as it is universally recognized as funding fundamental scientific research and education programs. We organize the funding opportunities into five major categories: (1) Early Career, (2) Infrastructure Oriented, (3) Core Research, (4) Transition to Practice, and (5) Education-Oriented. For each funding opportunity, we list its associated directorate, division, and funding range(s).

Early career funding aims to provide selected promising junior faculty with support to launch their research programs. Three major funding sources are available in this category: CRII, CAREER, and PECASE. CRII is targeted at junior faculty within the first 3 years of their career, while CAREER and PECASE are for junior faculty closer to their tenure stage. Infrastructure-oriented grants are commonly awarded out of the CISE directorate's Office of Advanced Cyberinfrastructure (OAC) and aim to help facilities support high-impact research. For example, DIBBs provide funding to build large-scale scientific testbeds to facilitate fundamentally sound research. CCRI

Table 3. Summary of Selected National Science Foundation (NSF) Funding Opportunities to Support AI for Cybersecurity Research and Education Programs

Funding Type	Selected Funding Opportunity	Directorate and Division	Funding Ranges
Early Career	CRII	CISE	Up to $175K
	CAREER	Cross-cutting	Up to $500K
	PECASE	Cross-cutting	Up to $500K
Infrastructure-Oriented	CSSI (formerly DIBBs)	CISE OAC	$200K–$1M
	CCRI	CISE OAC	Up to $1.2M
Core Research	SaTC CORE	Cross-cutting	$500K–$1.2M
	CICI	CISE OAC	$500K–$1M
	D-ISN	Cross-cutting	$250K–$1M
Transition to Practice	SaTC TTP	Cross-cutting	$500K–$1.2M
	SBIR/STTR	Cross-cutting	Up to $1.75M of seed funding
	Convergence Accelerator	Cross-cutting	$3M–$5M
Education-Oriented	SFS	EHR DGE	Varies
	SaTC-EDU	Cross-cutting	Up to $500K
	EAGER AI4Cyber	EHR DGE	Up to $300K

[*] *Note*: CCRI = CISE Community Research Infrastructure; CISE = Computer and Information Sciences and Engineering; CRII = CISE Research Initiation Initiative; CSSI = Cyberinfrastructure for Sustained Scientific Innovation; DGE = Division of Graduate Education; DIBBs = Data Infrastructure Building Blocks; D-ISN = Disrupting Illicit Supply Networks; EAGER = Early-Concept Grants for Exploratory Research; EHR = Education and Human Resources; OAC = Office of Advanced Cyberinfrastructure; PECASE = Presidential CAREER; SFS = Scholarship-for-Service; SaTC = Secure and Trustworthy Cyberspace; SaTC-EDU = SaTC Education; SBIR = Small Business Innovation Research Program; STTR = Small Business Technology Transfer Program; TTP = Transition to Practice.

provides funding to help scholars build computational infrastructure to facilitate their research. Both DIBBs and CCRI can provide valuable mechanisms to the rapidly growing AI for Cybersecurity community in terms of datasets and resources to support core research topics. Example programs within the NSF that fund cybersecurity-related research include SaTC, CICI, and D-ISN. SaTC focuses on fundamental cybersecurity-related research, particularly from an interdisciplinary perspective. CICI aims to fund research that aims to protect scientific cyberinfrastructure from attack, particularly by leveraging vulnerability assessment and netflow data. D-ISN is focused on disrupting illicit supply networks on the web by closely examining OSINT data sources (e.g., Dark Web forums, markets, etc.).

Oftentimes, research generated from the aforementioned grants may have significant practical utility. However, transitioning selected technologies to practice can often be a non-trivial task due to lack of a clear pathway, unclear end-users, and appropriate financial support. In recognition of these challenges, the NSF also has mechanisms to support faculty in commercializing their research and technologies to maximize their societal impact. Example programs to support this goal include SaTC TTP, SBIR/STTR, and Convergence Accelerator. Each opportunity emphasizes the critical importance of clear practical goals and impact. Finally, they provide knowledge about how AI for Cybersecurity (from scholarly and/or practical perspectives) can be delivered via innovative education programs. These programs can be funded by education-oriented funding, such as EAGER AI4Cyber and SaTC-EDU.

Each funding mechanism listed above can provide excellent resources to help scholars generate AI for Cybersecurity research at a rapid pace. However, peer-reviewing and archiving this research

Table 4. Selected Relevant Conferences and Venues for Disseminating AI for Cybersecurity Research

Conference Type	Selected Conference*	Approximate Annual Size	Relevant AI for Cybersecurity Workshop(s)
Academic Security Venues	IEEE S&P	700+	DLS
	ACM CCS	1,000+	AISec
	USENIX	2,000+	ScAINet
	IEEE ISI	200+	–
Industry Security Venues	Cyber Defense	1,000+	–
	DEFCON	20,000+	AI Village
	CAMLIS	100+	–
CS AI Venues	ACM KDD	3,000+	ISI-KDD
	ASONAM	1,000+	FOSINT-SI
	NeurIPS	13,000+	Trustworthy ML
	IEEE ICDM	800+	DL-CTI
NSF Meetings	SaTC PI Meeting	500+	AI for Cyber
	Trusted CI Meeting	500+	–
	SFS Job Fair	1,000+	–

* *Note*: AISec = Artificial Intelligence for Security; CAMLIS = Conference on Applied Machine Learning for Information Security; CCS = Computer and Communications Security; DL-CTI = Deep Learning for Cyber Threat Intelligence; DLS = Deep Learning for Security; FOSINT-SI = Foundations of Open Source Intelligence and Security Informatics; ISI = Intelligence and Security Informatics; ISI-KDD = Intelligence and Security Informatics Knowledge Discovery from Databases; KDD = Knowledge Discovery from Databases; ASONAM = Advances in Social Network Analysis and Mining; NeurIPS = Neural Information Processing Systems; ICDM = International Conference on Data Mining; SaTC = Secure and Trustworthy Cyberspace; S&P = Security and Privacy; SFS = Scholarship-for-Service; ScAINet = Security and AI Networking Summit.

is essential for ensuring the long-term viability and health of a discipline. Conferences and journals can provide excellent mechanisms in this regard. We summarize a selected listing of prevailing cybersecurity meetings and venues in Table 4. We group the opportunities into four major groups: academic security venues, industry security venues, computer science AI venues, and NSF meetings. For each conference, we also summarize its approximate annual size as well as relevant AI for cybersecurity workshops.

Each conference venue offers a tremendous opportunity for AI for Cybersecurity scholars and practitioners to share and disseminate ideas as well as network with colleagues. For example, academic venues are increasingly supporting workshops pertaining to various topics related to CTI, adversarial ML, trustworthy computing, and other AI for Cybersecurity topics. These conferences can also serve as an excellent mechanism to get feedback on preliminary work for possible extension to prevailing cybersecurity journals and magazines, including *IEEE Transactions on Dependable and Secure Computing* (TDSC), *IEEE Transactions on Information Forensics and Security* (TIFS), *ACM Transactions on Privacy and Security* (TOPS; formerly TISSEC), *IEEE Security and Privacy Magazine*, and *Computers and Security* (C&S). Published works in these journals can help archive the research and contribute to the long-term viability of the AI for Cybersecurity discipline.

6 ARTICLES IN THIS SPECIAL ISSUE

This special issue aims to take an important step in cultivating the AI for Cybersecurity community. The idea for this special issue was originally conceived at the International Conference on Information Systems (ICIS) in December 2018 in San Francisco, California. In the ensuing months, the guest editor team was developed and the Call for Papers (CFP) was carefully refined. The initial

Table 5. Summary of Articles in This Special Issue

Authors	Title	Topic	Data Sources	Methodology
Husak et al.	Predictive Cyber Situational Awareness and Personalized Black Listing: A Sequential Rule Mining Approach	Personalized blacklisting	12 million alerts from 34 IDs	Sequential rule mining
Mangino et al.	Internet-scale Insecurity of Consumer Internet of Things: An Empirical Measurements Perspective	Fingerprinting infected IoT devices	3.6TB of network traffic data; 800K compromised IPs	Gaussian Naïve Bayes, SVM, Random Forest, intelligent feature selection
Mudgerikar et al.	Edge-based Intrusion Detection for IoT Devices	Profiling IoT devices	3,973 traditional IoT malware samples	Random Forest and Naïve Bayes
Namayanja et al.	IP Reputation Scoring with Geo-Contextual Feature Augmentation	Effective anomaly detection for encrypted network data based on geo-location reputations	See blacklisted and whitelisted data, network data sources, geo-contextual data	Geo-spatial analysis, clustering, reputation services
Shao et al.	An Ensemble of Ensembles Approach to Author Attribution for Internet Relay Chat Forensics	Author identification and threat message detection for IRC monitoring	Nine IRC channels pertaining to hacking activities	Deep forest and ensemble methodologies
Sharma et al.	Panda: Partitioned Data Security on Outsourced Sensitive and Non-Sensitive Data	Encrypting sensitive vs. non-sensitive outsourced data	MPC-based Jana and SGX-based Opaque	Query binning; cryptographic techniques
Sweet et al.	On the Variety and Veracity of Cyber Intrusion Alerts Synthesized by Generative Adversarial Networks	Generating synthetic alerts to emulate critical dependencies	Alert logs	Generative adversarial networks
Zhang et al.	Analysis of Cyber Incident Categories Based on Losses	Quantifying risks for insurance and risk management applications	VERIS Database, WebHacking Incident Database, Privacy Rights Clearinghouse	Loss occurrences with Bernoulli random variables, clustering, network visualizations
Zihayat et al.	A Time-based Gap Analysis of Cybersecurity Trends in Academic and Digital Media	Identifying emerging cybersecurity trends between 2008 and 2018	3,556 academic papers and 4,163 *New York Times* articles	Topic modeling, temporal analysis

CFP was distributed via the AISWorld listserv; DBWorld listserv; the 2019 IEEE Intelligence and Security Informatics (ISI) Conference in Shenzhen, China; DEFCON AI Village in Las Vegas; and over 250 personal emails in our contact networks. These efforts resulted in 60 submissions to the special issue. Sixteen papers were desk rejected due to lack of fit, resulting in 44 total papers being sent out for review. During the peer review process, an additional 29 papers were rejected. Out of the final accepted papers, nine were eventually selected to be part of the cybersecurity related special issue. Each paper went through two or three rounds of review. We provide a summary of the final accepted papers in Table 5. For each paper, we identify the topic, data sources, and methodology employed or developed. Selected impacts the work has on AI for Cybersecurity are noted following the table. Papers are listed in alphabetical order based on the last name of the first author.

Each of the articles made interesting and timely contributions to the field of AI for Cybersecurity. In the article entitled "Predictive Cyber Situational Awareness and Personalized Black Listing: A Sequential Rule Mining Approach," Husak et al. employed a sequential rule mining approach on 12 million alerts from 34 Intrusion Detection Systems for personal blacklisting applications. Results of

this work can significantly help SOC analysts in reducing alert fatigue. Mangino et al. contributed an article entitled "Internet-scale Insecurity of Consumer Internet of Things: An Empirical Measurements Perspective" that focused on fingerprinting IoT-infected IP addresses via an intelligent feature selection and classification algorithm approach. Results of their work can have significant implications for facilitating global CTI regarding the profile of malicious IoT devices. Mudgerikar et al. aimed to study a related phenomenon of profiling IoT devices in their work entitled "Edge-based Intrusion Detection for IoT Devices." In particular, they leveraged a classification methodology on 3,973 malware samples to achieve their goal. Namayanja et al. aimed to gain a global understanding of device reputations for anomaly detection applications in their article entitled "IP Reputation Scoring with Geo-Contextual Feature Augmentation." Similar to Mangino et al., this work has significant implications for mapping out the global threat landscape. Shao et al. aimed to also make strides in the CTI area in their work entitled "An Ensemble of Ensembles Approach to Author Attribution for Internet Relay Chat Forensics." In particular, they analyzed nine IRC channels via a novel author attribution deep forest approach with the goal of attributing selected authors in IRC channels and detecting selected threat messages. Sharma et al. aimed to tackle the problem of anonymizing sensitive data via a query binning and cryptographic techniques in their article entitled "Panda: Partitioned Data Security on Outsourced Sensitive and Non-Sensitive Data." Sweet et al. contributed an article entitled "On the Variety and Veracity of Cyber Intrusion Alerts Synthesized by Generative Adversarial Networks." This work makes excellent contributions to SOC analysts and adversarial ML by tackling the key issue of alert analysis. In the article "Analysis of Cyber Incident Categories Based on Losses," by Zhang et al. leveraged Bernoulli random variables, clustering, and network analytics to quantify the risk of selected cyber-attacks for insurance and risk management applications. Their work has significant implications for strategic cybersecurity professionals who focus on making critical cybersecurity investments. Finally, Zihayat et al. aimed to identify emerging trends between 2008 and 2018 via a topic modeling and temporal analysis-based approach in their work entitled "A Time-based Gap Analysis of Cybersecurity Trends in Academic and Digital Media." Their work provides a powerful approach for keeping up with the ever-evolving cybersecurity landscape.

7 SUMMARY

Preventing cyber-attacks has become a grand societal challenge. Despite significant investments in various cybersecurity programs, breaches remain on an upward trend. AI-based techniques hold significant promise in sifting through large quantities of heterogeneous cybersecurity data to efficiently and effectively support critical cybersecurity tasks such as asset prioritization, controls allocation, threat detection, and vulnerability management. Despite these potential benefits, the AI for Cybersecurity discipline is still in its nascency. Numerous opportunities exist for scholars to make significant progress and practical impacts in turning the tide against cyber-attacks.

In this article, we aimed to provide an important step to progressing the AI for Cybersecurity discipline. In particular, we provided an overview of prevailing cybersecurity data (categorized into internal and external sources), summarized extant application areas of cybersecurity, and identified key limitations within the prevailing landscape. Based on these key issues, we offered a multi-disciplinary AI for Cybersecurity roadmap that centers on major themes such as cybersecurity applications and data, advanced AI methodologies for cybersecurity, and AI-enabled decision making. To help scholars and practitioners make significant headway in tackling these grand AI for Cybersecurity issues, we summarized promising funding mechanisms that can support long-term, systematic research programs and summarized prevailing journal and conference venues for disseminating AI for Cybersecurity research. We hope that the contents of this article, along

with the articles in this special issue, stimulate discussion that can lead to the rapid growth of high-impact AI for Cybersecurity topics in a manner that positively impacts society.

ACKNOWLEDGMENTS

We thank the numerous reviewers for offering their valuable time to provide excellent feedback to the submitted article. We thank Editor-in-Chief Dr. Daniel Zeng for his guidance and feedback throughout the special issue process. We thank Victoria White for her coordination of papers, reviews, and feedback.

REFERENCES

[1] Ahmed Abbasi, Suprateek Sarker, and Roger Chiang. 2016. Big data research in information systems: Toward an inclusive research agenda. *J. Assoc. Inf. Syst.* 17, 2 (2016), I–XXXII. DOI: https://doi.org/10.17705/1jais.00423

[2] Ahmed Abbasi, Zhu Zhang, David Zimbra, Hsinchun Chen, and Jay F Nunamaker. 2010. Detecting fake websites: The contribution of statistical learning theory. *MIS Q.* (2010), 435–461.

[3] Nolan Arnold, Mohammadreza Ebrahimi, Ning Zhang, Ben Lazarine, Mark Patton, Hsinchun Chen, and Sagar Samtani. 2019. Dark-net ecosystem cyber-threat intelligence (CTI) tool. In *2019 IEEE International Conference on Intelligence and Security Informatics (ISI'19)*. DOI: https://doi.org/10.1109/ISI.2019.8823501

[4] Indranil Bardhan, Hsinchun Chen, and Elena Karahanna. 2020. Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management. *MIS Q.* 44, 1 (2020), 185–200.

[5] Katy Börner and David E. Polley. 2014. *Visual Insights: A Practical Guide to Making Sense of Data*. MIT Press.

[6] Matt Bromiley. 2016. Threat intelligence: What it is, and how to use it effectively. *SANS Institute*. Retrieved June 5, 2017, from https://www.sans.org/reading-room/whitepapers/analyst/threat-intelligence-is-effectively-37282.

[7] Miles Brundage, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitzoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt, Carrick Flynn, Seán Ó HÉigeartaigh, Simon Beard, Haydn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crootof, Owain Evans, Michael Page, Joanna Bryson, Roman Yampolskiy, and Dario Amodei. 2018. The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. (February 2018). Retrieved from http://arxiv.org/abs/1802.07228

[8] Hsinchun Chen. 2012. *Dark Web: Exploring and Data Mining the Dark Side of the Web*. Springer New York, New York, NY. DOI: https://doi.org/10.1007/978-1-4614-1557-2

[9] Hsinchun Chen, Roger H. L. Chiang, and Veda C. Storey. 2012. Business intelligence and analytics: From big data to big impact. *MIS Q.* 36, 4 (2012), 1165–1188. DOI: https://doi.org/10.1145/2463676.2463712

[10] Ronan Collobert and Jason Weston. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th International Conference on Machine learning (ICML'08)*, 160–167. DOI: https://doi.org/10.1145/1390156.1390177

[11] Mengnan Du, Ninghao Liu, and Xia Hu. 2019. Techniques for interpretable machine learning. *Commun. ACM* 63, 1 (2019), 68–77. DOI: https://doi.org/10.1145/3359786

[12] Po-Yi Du, Ning Zhang, Mohammedreza Ebrahimi, Sagar Samtani, Ben Lazarine, Nolan Arnold, Rachael Dunn, Sandeep Suntwal, Guadalupe Angeles, Robert Schweitzer, and Hsinchun Chen. 2018. Identifying, collecting, and presenting hacker community data: Forums, IRC, carding shops, and DNMs. In *2018 IEEE International Conference on Intelligence and Security Informatics (ISI'18)*, 70–75. DOI: https://doi.org/10.1109/ISI.2018.8587327

[13] Malaka El, Emma McMahon, Sagar Samtani, Mark Patton, and Hsinchun Chen. 2017. Benchmarking vulnerability scanners: An experiment on SCADA devices and scientific Instruments. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI'17)*, 83–88. DOI: https://doi.org/10.1109/ISI.2017.8004879

[14] Katheryn A. Farris, Ankit Shah, George Cybenko, Rajesh Ganesan, and Sushil Jajodia. 2018. VULCON: A system for vulnerability prioritization, mitigation, and management. *ACM Trans. Priv. Secur.* 21, 4 (2018), 1–28. DOI: https://doi.org/10.1145/3196884

[15] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in Neural Information Processing Systems*.

[16] Shirley Gregor and Alan R. Hevner. 2013. Positioning and presenting design science research for maximum impact. *MIS Q.* 37, 2 (2013), 337–355. DOI: https://doi.org/10.2753/MIS0742-1222240302

[17] John Grisham, Sagar Samtani, Mark Patton, and Hsinchun Chen. 2017. Identifying mobile malware and key threat actors in online hacker forums for proactive cyber threat intelligence. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI'17)*, 13–18. DOI: https://doi.org/10.1109/ISI.2017.8004867

[18] Christopher R. Harrell, Mark Patton, Hsinchun Chen, and Sagar Samtani. 2018. Vulnerability assessment, remediation, and automated reporting: Case studies of higher education institutions. In *2018 IEEE International Conference on Intelligence and Security Informatics (ISI'18)*. DOI: https://doi.org/10.1109/ISI.2018.8587380

[19] Hemant Jain, Balaji Padmanabhan, Paul A. Pavlou, and Raghu T. Santanam. 2018. Humans, algorithms, and augmented intelligence: The future of work, organizations, and society. *Inf. Syst. Res.* 29, 1 (2018), 250–251. DOI: https://doi.org/10.1287/isre.2018.0784

[20] Anne Johnson and Emily Grumbling (Eds.). 2019. *Implications of Artificial Intelligence for Cybersecurity*. National Academies Press, Washington, DC. DOI: https://doi.org/10.17226/25488

[21] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature* 521, 7553 (2015), 436–444. DOI: https://doi.org/10.1038/nature14539

[22] Qing Li and Gregory Clark. 2015. *Security Intelligence: A Practitioner's Guide to Solving Enterprise Security Challenges*. John Wiley & Sons, Inc., Indianapolis, IN.

[23] Yunji Liang, Sagar Samtani, Bin Guo, and Zhiwen Yu. 2020. Behavioral biometrics for continuous authentication in the internet-of-things era: An artificial intelligence perspective. *IEEE Internet Things J.* 7, 9 (2020), 9128–9143. DOI: https://doi.org/10.1109/JIOT.2020.3004077

[24] Raffael Marty. 2008. *Applied Security Visualization*. Addison-Wesley Professional.

[25] National Science and Technology Council. 2019. *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update*. Washington, DC. Retrieved from https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf.

[26] National Science Foundation. 2019. National artificial intelligence (AI) research institutes (2019). nsf20503 | NSF – national science. Retrieved from https://www.nsf.gov/pubs/2020/nsf20503/nsf20503.pdf.

[27] Jay F. Nunamaker, Nathan W. Twyman, Justin Scott Giboney, and Robert O. Briggs. 2017. Creating high-value real-world impact through systematic programs of research. *MIS Q.* 41, 2 (2017), 335–351. DOI: https://doi.org/10.25300/MISQ/2017/41.2.01

[28] Jay F. Nunamaker, Minder Chen, and Titus D. M. Purdin. 1990. Systems development in information systems research. *J. Manag. Inf. Syst.* 7, 3 (1990), 89–106.

[29] Alessandro Parisi. 2019. *Hands-On Artificial Intelligence for Cybersecurity: Implement Smart AI Systems for Preventing Cyber Attacks and Detecting Threats and Network Anomalies*. Packt Publishing, Birmingham, UK.

[30] Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee. 2007. A design science research methodology for information systems research. *J. Manag. Inf. Syst.* 24, 3 (2007), 45–77.

[31] Arun Rai. 2020. Explainable AI: From black box to glass box. *J. Acad. Mark. Sci.* 48, 1 (2020), 137–141. DOI: https://doi.org/10.1007/s11747-019-00710-5

[32] Arun Rai, Panos Constantinides, and Saonee Sarker. 2018. Editor's comments: Next-generation digital platforms: Toward human–AI hybrids. *MIS Q.* 43, 1 (2018), iii–ix.

[33] S. Samtani, H. Zhu, and H. Chen. 2020. Proactively identifying emerging hacker threats from the dark web. *ACM Trans. Priv. Secur.* 23, 4 (2020), 1–33. DOI: https://doi.org/10.1145/3409289

[34] Sagar Samtani, Maggie Abate, Victor Benjamin, and Weifeng Li. 2020. Cybersecurity as an industry: A cyber threat intelligence perspective. In *The Palgrave Handbook of International Cybercrime and Cyberdeviance*. DOI: https://doi.org/10.1007/978-3-319-78440-3_8

[35] Sagar Samtani, Kory Chinn, Cathy Larson, and Hsinchun Chen. 2016. Azsecure hacker assets portal: Cyber threat intelligence and malware analysis. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI'16)*, 19–24. DOI: https://doi.org/10.1109/ISI.2016.7745437

[36] Sagar Samtani, Ryan Chinn, and Hsinchun Chen. 2015. Exploring hacker assets in underground forums. In *2015 IEEE International Conference on Intelligence and Security Informatics (ISI'15)*, 31–36. DOI: https://doi.org/10.1109/ISI.2015.7165935

[37] Sagar Samtani, Ryan Chinn, Hsinchun Chen, and Jay F. Nunamaker. 2017. Exploring emerging hacker assets and key hackers for proactive cyber threat intelligence. *J. Manag. Inf. Syst.* 34, 4 (2017), 1023–1053.

[38] Sagar Samtani, Shuo Yu, Hongyi Zhu, Mark Patton, and Hsinchun Chen. 2016. Identifying SCADA vulnerabilities using passive and active vulnerability assessment techniques. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI'16)*, 25–30. DOI: https://doi.org/10.1109/ISI.2016.7745438

[39] Sagar Samtani, Hongyi Zhu, Balaji Padmanabhan, Yidong Chai, and Hsinchun Chen. 2020. Deep learning for information systems research. (October 2020). Retrieved from http://arxiv.org/abs/2010.05774

[40] Joshua Saxe and Hillary Sanders. 2018. *Malware Data Science: Attack Detection and Attribution*. No Starch Press, San Francisco, CA.

[41] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. 2016. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (6th ed.). Pearson.

[42] Mark Stamp. 2017. *Introduction to Machine Learning with Applications in Information Security*. CRC Press, Taylor & Francis Group, Boca Raton, FL.

[43] Yonghui Xu, Sinno Jialin Pan, Hui Xiong, Qingyao Wu, Ronghua Luo, Huaqing Min, and Hengjie Song. 2017. A unified framework for metric transfer learning. *IEEE Trans. Knowl. Data Eng.* 29, 6 (2017), 1158–1171. DOI: https://doi.org/10.1109/TKDE.2017.2669193

[44] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning. *ACM Trans. Intell. Syst. Technol.* 10, 2 (2019), 1–19. DOI: https://doi.org/10.1145/3298981

[45] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. 2014. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems*.

[46] Hongyi Zhu, Sagar Samtani, Randall Brown, and Hsinchun Chen. 2020. A Deep Learning approach for recognizing activity of daily living (ADL) for senior care: Exploiting interaction dependency and temporal patterns. *MIS Q.* (2020), Forthcoming. Retrieved from https://ssrn.com/abstract=3595738.

[47] Hongyi Zhu, Sagar Samtani, Hsinchun Chen, and Jay F. Nunamaker. 2020. Human identification for activities of daily living: A deep transfer learning approach. *J. Manag. Inf. Syst.* 37, 2 (2020), 457–483. DOI: https://doi.org/10.1080/07421222.2020.1759961

Received October 2020; revised November 2020; accepted November 2020