



Transparency as Threat at the Intersection of Artificial Intelligence and Cyberbiosecurity

Sara R. Jordan and Samantha L. Fenn, Virginia Polytechnic Institute and State University
Benjamin B. Shannon, Federal Bureau of Investigation

Have separate actions designed to meet ethical norms of transparency inadvertently led to a situation of greater possible harm from biological warfare attacks? Efforts to make artificial intelligence and data from biological sciences more publicly available have raised novel concerns about national security.

What if the threat posed by artificial intelligence (AI) will come from microscopic agents built using AI and not from anthropomorphic robots that directly attack humans? What if the necessary components to build that attack arise from the combination of biological and programming components freely available as a consequence of humans striving to be ethical? The literature surrounding both physical biosecurity (to include recombinant DNA and synthetic biology) and AI (to

include autonomous systems) is leavened with warnings of potential catastrophe if risks are not well managed.

As a package of tools without independent or inherent substantive commitments, the risks from AI to humans, including risks to national security, are domain specific. When AI tools are deployed using materials from a domain whose characteristics would already raise concerns when a human intelligence is applied, there are elevated concerns about the use of AI. In the context of biological data, the use of AI is one area of elevated threat. In brief, openly sourced, faster, more efficient, and scalable algorithms could be placed against openly available, publicly accessible biological databases to more easily,

Digital Object Identifier 10.1109/MC.2020.2995578
 Date of current version: 5 October 2020

quickly, often, and cheaply identify agents or their components with lethal or harmful effects, than if AI tools were not used. As such, AI stands to scale and expedite the problems of dual-use research (DUR) in biosecurity and, particularly, the cyberbiosecurity domain at relatively low cost and potentially out of sight of conventional tracking systems.

Of specific concern are the potential uses of data from biological pathogens regulated by the Federal Select Agent Program that pose threats to human safety and national security and, as a result, have historically been tightly monitored in highly controlled environments. The national security threat that arises from AI and sensitive biological data comes from the synergy of the two. When biological threat actors (humans) exploiting nonhuman biotic systems (plants, animals, marine environments, and so on) interact with the cyberbiosecurity domain, which makes data from bacteriological, virological, toxicological, fungal, or parasitic infectious agents machine readable and interpretable using an Internet connection, the new threat emerges.

The introduction of computer-aided data mining, massively open access databases, and decision-scoring systems into the cyberbiosecurity domain raises the possibility that the tools of AI could be used to identify and locate agents of concern, or even components of agents, that could then be incorporated into biological systems to amplify pathogenicity or lethality. These same tools could also be used to confound efforts to develop countermeasures by maliciously redirecting the training of “white hat” algorithms that search for useful or lethal agents or their countermeasures.

We take the position that the transparency of AI methods and biological data presents a novel case for national security. Specifically, the drive for ethical behavior in both fields has resulted in a situation that is understood as having an unacceptable probability of leading to harm without regulation, including the establishment of a network of (human) tripwires in the public health and computing communities that would create open lines of communication between the scientific and the intelligence communities. The intersection of open source AI and open source biological data leads to amplified worries regarding dual-use research of concern (DURC) in both areas. DUR is research conducted for legitimate purposes that generates knowledge, information, technologies, and/or products that can be utilized for both benevolent and harmful purposes.

We unfold this argument as follows. First, we identify how the ethics of science may have led to this scenario. Second, we review the dual-use concerns related to AI and biological data. Third, we propose mechanisms that could mitigate the dual-use threats for the intersection of the two domains and explain the challenges these pose in terms of the ethics of accountability and transparency for both. Fourth, we propose a modified ethic of transparency and accountability for both AI and cyberbiosecurity so that risk-facing situations may not inadvertently arise.

CYBERBIOSECURITY

Until recently, the biodefense and biosecurity conversation has revolved primarily around the physical storage and handling of biological agents that pose threats for bioterrorism attacks and the potential to start a pandemic,

either by an accidental or intentional release of a dangerous microbiological pathogen. As a result of advances in cybercapabilities to include publicly accessible biological data warehouses, a new concern, labeled cyberbiosecurity, emerged (Figure 1). Cyberbiosecurity

aims to direct attention to the potential of cyberattacks into biological processes that would have major reputational, economic, and national security ramifications for many sectors, such as health-care, biopharmaceutical, agricultural, and genomics, among other biomanufacturing processes. As the new discipline is still being defined, it is becoming apparent that there are possibilities of cyber intrusion into biological processes of many biological-based practices. Cyberbiosecurity is being proposed as a formal new enterprise, which encompasses cybersecurity, cyberphysical security and biosecurity as applied to biological and biomedical-based systems.¹

Cyberbiosecurity experts advocate for a forward-thinking mindset for national security, which emphasizes the importance of

understanding the vulnerabilities to unwanted surveillance, intrusions, and malicious and harmful activities, which can occur within or at the interfaces of commingled life and medical sciences, cyber, cyberphysical, supply chain and infrastructure systems, and developing and instituting measures to prevent, protect against, mitigate, investigate and attribute such threats as it pertains to security, competitiveness and resilience.¹

TRANSPARENCY AND NOVEL SECURITY THREATS

AI and biosecurity concerns were brought together as a consequence of the serendipity of two scientific domains seeking to achieve the same ethical goals: transparency and accountability. Within the many discussions of ethics for scientific practice, transparency of scientific methodology, material (for example, data), and funding stands out as essential for good publications, fundable grant projects, and the accumulation of knowledge within the disciplines. To the extent that methods, materials, and funding are made transparent to readers, the scientific community is understood to have fulfilled its basic obligations to the public and to the ethical value of transparency.

For AI, the drive for accountability and transparency relates to methods of inquiry and deployment. As will be described more fully later in this article, when these methods are showcased through public demonstrations—open source code on professional databases (for example, GitHub), explanatory videos, peer-reviewed publications, popular publications, and massively open online courses—transparency is more fully achieved. However, when such transparency is achieved, the otherwise neutral methods can be applied to data of virtually any variety, including biological data.

For biosecurity, the drive for accountability and transparency relates to the materials or data resulting from inquiry. Public biological databases open to the scientific community house an ample amount of information that can be used to treat infectious diseases, locate organisms that can lead to outbreaks or pandemics, and identify cellular mechanisms to gain a better

understanding of illnesses with no current treatment or cure.³

The situation of AI and cyberbiosecurity is one in which the drive for transparency and accountability of methods and materials may lead to situations where research data could be used in either positive or negative ways. The concern that data could be put to such dual use arises when a malicious actor has access to data or uses them to exploit vulnerabilities. The amplification of these concerns arises from the ready availability of both AI methods and biological data (material) and methods (for example, plasmid transformations) on the Internet, including their proliferation as a commodity on the “dark web”; what was formerly the province of a few experts in control of the data or the code is now open source property of the public at large.⁴

Transparency as an ethical norm for science

Protecting the lay community while advancing the ends of science is a

balancing act familiar to many scientists.⁵ Wrapped in many guises, whether statements of scientific ethics or processes for the ethical review of proposed research, the balance of scientific progress and national security is foundational to physical, social, and computational scientific work. One way in which this balance manifests is in the drive for publicity in science. Given arguments that the public has a “right to know” what research is conducted in their communities, scientists have responded with a number of initiatives that increase transparency and knowledge among the lay population and also increase the transparency of data for science.⁶ However, as experts in the national security domain suggest, a level of unacceptable transparency in science is reached when national security concerns arise.

According to major research ethics documents, the ethical values of science include honesty, fairness, objectivity, openness, trustworthiness, and

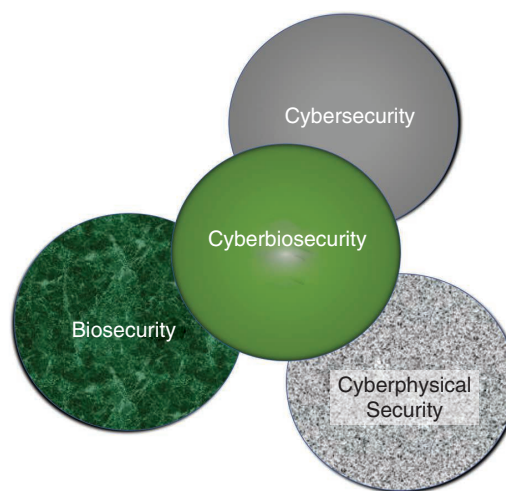


FIGURE 1. A description of cyberbiosecurity as an emerging discipline that covers the cybersecurity, biosecurity, and cyberphysical security aspects of biological processes.²

respect for others. Resnick characterized “openness” as willingness to “share data, results, ideas, methods, and techniques with other researchers.”⁷ The Nuclear Regulatory Commission suggested that the benefits of data sharing include “maintaining the integrity of the research process by providing independent opportunities for verification, refutation, or refinement of original results and data; promoting new research and the development and testing of new theories; and encouraging appropriate use of empirical data in policy formulation and evaluation.”⁸ Within the global community of science, broad pronouncements about the value of research data sharing, such as is found in the Singapore Statement on Research Integrity, encourage researchers to “share data and findings openly and promptly as soon as they have had an opportunity to establish priority and ownership claims.”⁹ This and other general statements also ask researchers to “weigh societal benefits against risks inherent in their work” but do not address, specifically, the role of research in promoting or diminishing national security.

Ethics in AI and computing

The Association for Computing Machinery (ACM) updated its code of ethics and professional conduct in 2018, affirming that “computing professionals’ actions change the world. To act responsibly, they should reflect upon the wider impacts of their work, consistently supporting the public good.”¹⁰ According to recent research, the ACM code is but one of many major guidelines set to govern the ethics of computer science, including AI research and product development. While there are myriad principles proposed to govern AI, Floridi and Cows

argue against this “principle proliferation” and suggest that there are five major principles: “beneficence, non-maleficence, autonomy, justice, and explicability (which is divided into intelligibility and accountability).”¹¹

Many statements emphasize the principles of accountability, transparency, or explainability. Explainable AI, or the effort to open the “black box” of sophisticated AI techniques like neural networks,

is a multifaceted topic. It encompasses both individual models and the larger systems that incorporate them. It refers not only to whether the decisions a model outputs are interpretable, but also whether or not the whole process and intention surrounding the model can be properly accounted for.

There need to be three steps which should be fulfilled by the system:

1. *explain the intent behind how the system affects the concerned parties*
2. *explain the data sources you use and how you audit outcomes*
3. *explain how inputs in a model lead to outputs.*¹²

The development of explainable AI to include methods for explaining the data and the models or to unpack biased conclusions, misclassification, or behavior that contradicts developers’ or users’ expectations, explaining both the materials (data) and methods (models), motivates conversations about ethical AI.

Ethics for biological data

While traditional concerns of informed consent, clinical equipoise, and the

accumulation of valid scientific knowledge continue to motivate research ethics work, the advent of clinical big data and telemedicine added important elements to these discussions.¹³ The formal definition of *biological big data* is “massive, diverse data sets that are created, reside, are analyzed and move in information ecosystems. For the life sciences Big Data refers to datasets including ‘raw data, combined data, or published data from the health-care system, pharmaceutical industry, genomics and other -omics fields, clinical research, environment, agriculture, and microbiome efforts.’”¹ Biological big data also include “analytic technologies and outputs, such as from data integration, data mining, data fusion, image and speech recognition, natural language processing, machine learning, social media analysis, and Bayesian analysis.”¹ Biological big data are sometimes analyzed in conjunction with various formats of biological data, such as that from genealogy services (for example, 23&Me, AncestryDNA) that offer users insight into heritage, genetic traits, and ancestral history.

Biological data, which include biomedical data, can be used for many applications, both beneficent and malevolent, which confounds the breadth of application of dual-use controls. Microlevel personal data can be used to improve clinical outcomes in oncology, but these same data can also be used by nefarious actors, such as when medical history from an individual is stolen. Health-care fraud may lead to data theft and corporate espionage or can be used against individuals for blackmail. Experts in cyberbiosecurity have expressed major concerns with health-care data exploitation, such as through the

manipulation of personalized medical practices.

U.S. Federal Bureau of Investigation (FBI) Supervisory Special Agent (SSA) Edward You, of the Weapons of Mass Destruction Directorate's (WMDD's) Biological Countermeasures Unit, is one of the subject matter experts in the dual-use concern of health-care data. SSA You expressed the following concern:

If a foreign source, especially a criminal one, has your biological information, then they might have some particular insights into what your future medical needs might be and exploit that. For instance, what happens if you have a singular medical condition and an outside entity says they have a treatment for your condition? You could get talked into paying a huge sum of money for a treatment that ends up being bogus.¹⁴

In a larger-impact incident, an intrusion into personalized biological big data could be used to exploit groups within government agencies and gain access to data warehouses with the intent of stealing government or personal intellectual property.

The misuse of biological data, whether health-care data, medical records, or genomic information, can have security repercussions for both individuals and groups as large as nations. For example, in conjunction with the FBI's WMDD, the National Academies of Sciences, Engineering, and Medicine hosted a workshop for "understanding the applications and potential security implications of emerging technologies at the interface of the life sciences and information sciences."¹⁵

One speaker described these issues as including

... phishing scams, bots, and ransomware, to software holes that leave vulnerabilities, with particular attention to the rise of phishing scams and ransomware used to target health-care data. He noted that there has been a 3,500% increase in the use of ransomware attacks, leading to the payment of more than US\$1.6 million, yet these figures are potentially underreported, as there is no mandate for private companies to disclose breaches to their consumers within the bioeconomy as there is in other sectors. Brooks discussed how the life science community is particularly vulnerable to cyberattacks because much of the data generation and analysis tools require digitalization and use of devices for which there are not yet adequate standards of protection.¹⁵

IDENTIFYING AND MITIGATING CYBERBIODATA THREATS

The threat of malicious agents creating novel biological agents through the use of open biological data and open AI is a real national security concern. Whether the agents developed represent adaptations to organisms that improve the lethality of existing agents, such as gain of function (GOF), or represent the development of agents from research designed to launch a counterstrategy to evolving medical countermeasures, there is potential that the use of AI tools could expedite or augment the nascent threat. The ability to induce a GOF property by manipulating the biological mechanisms within organisms

to alter the dangerousness of a pathogen is a dual-use technique that needs greater attention from the national security community. In this section, we identify two scenarios in which the intersection of AI and cyberbio data raises risks.

Two threat scenarios

Espionage and counter-countermeasures. Using information publicly available on <https://clinicaltrials.gov>, individuals from Company Z, a gene-sequencing equipment vendor, are able to identify that both University A and University B are recipients of grants to engage in research to develop countermeasures for a lethal pathogen. Both University A and University B utilize sequencing equipment from Company Z. Company Z grooms a graduate student, who is struggling financially, and offers significant compensation if the student provides access to the data at University A. The intention of Company Z is to gather all of University A's stored data to sell it to a buyer at University B who is interested in advancing his research speed and to sell it to other actors who express interest in countering countermeasures. Company Z offers to University A opportunities for webinars designed to help University A prepare its laboratory for smooth annual preventative maintenance of their equipment.

The graduate student is used as a contact to set up and enable "smooth webinars" within the laboratories of University A. The graduate student accesses the cloud storage infrastructure and uploads software allowing Company Z to send to each of the webinar participants a malware program that was made more harmful by the incorporation of a learning algorithm

to “learn” to evade filters. Once this malware is in the systems of University A, it allows buyers from University B and Company Z to have access to all data that are generated and stored in its cloud-based platforms. University B is able to “scoop” the findings of University A, leveraging its publications for further grants and contracts, and Company Z is able to sell the additional data to other parties through dark web channels.

Malicious intrusions and targeted health advertising.

Eager to move forward quickly to comply with mandates for electronic medical records issued by the Centers for Medicare and Medicaid Services, a medical office management company pursues vendors to move all of its data, en masse, to a cloud platform. The group chooses a vendor who offers to expedite the transition from paper documents to electronic data using an in-built suite of natural language processing options to identify, transform, and extract text. Due to the nature of the firms’ existing records and insufficient training on the separation of administrative or billing information from clients’ health information, patient health data and patient economic data are uploaded and extracted. Likewise, due to overeager purchasing decisions, the firm did not acknowledge that its information would be extracted by a subcontractor to the third-party vendor whose primary location of business is in a nation known to perpetrate online fraud on a large scale.

With the combination of patients’ health and personal information, the subcontractor is able to develop a targeted spear-phishing campaign against the medical offices’ clients. Using the records of individuals with

significant diseases and the firms’ own spoofed email addresses, fraudulent claims for precision medicine practices are sent to vulnerable patients. If a patient responds, the malicious actors may validate sensitive medical information and then receive access to the patient’s payments, possibly allowing the actors to drain the patient’s bank account with little to no traceability for identifying the perpetrator.

EVALUATING AI AND CYBERBIOSECURITY THREATS

How do we know how significant the threat from cases like this might be? Published in 1982, the “Corson report” outlined methods for defining when research is of sufficient concern about its dual use that it should fall into a “gray area” of research requiring regulation, such as is done through import–export control mechanisms. Within the Corson report, the panel of experts convened recommended the following:

No restrictions of any kind limiting access or communication should be applied to any area of university research, be it basic or applied, unless it involves a technology meeting all of the following criteria: 1) The technology is developing rapidly and the time from basic science to application is short; 2) the technology has identifiable direct military applications; or it is dual-use and involves process or production-related techniques; 3) transfer of the technology would give the USSR a significant near term military advantage; and the United States is the only source of information about the technology

or other friendly nations that could also be the source have control systems as secure as ours [numbering added].¹⁶

The changes to the research and political landscape in the 38 years since the Corson report have meant modifications to the four conditions, such as was done in 2017 by the Committee on DURC.¹⁷ Of significant importance was the 2011 controversy regarding the publication of GOF H5N1 influenza research.⁵ This controversy arose from the publication of “gain of function research of concern,” including the “production of H5N1 influenza A viruses that are airborne-transmissible among ferrets, compared to the nonairborne transmissible wild type.”¹⁸ The fact that this research became part of the open source environment of science created significant concerns about the weaponization of biological agents by nations or nefarious nonnational actors. Based upon experience with this controversy, the Corson report points were expanded to five key questions:

- ▶ Are there reasonably anticipated risks to public health and safety from direct misapplication of this information, that is, is novel scientific information provided that could be intentionally misused to threaten public health or safety?
- ▶ Are there reasonably anticipated risks to public health and safety from direct misapplication of this information, that is, does the information point out a vulnerability in public health and/or safety preparedness?
- ▶ Is it reasonably anticipated that this information could be directly misused to pose a

threat to agriculture, plants, animals, the environment, or materiel?

- › If a risk has been identified, in what timeframe (for example, immediate, near future, years from now) might this information be used to pose a threat to public health and/or safety, agriculture, plants, animals, the environment, or materiel?
- › If the information were to be broadly communicated “as is,” what is the potential for public misunderstanding, that is, what might be the implications of such misunderstandings (for example, psychological, social, health/dietary decisions, economic, commercial, etc.)? For sensationalism?¹⁶

In this report, it was pointed out that these questions covered only the realm of research that had an affiliate history with biosafety and biosecurity oversight, otherwise neglecting research that is “a theoretical modeling study,” which involves “big data and data mining tools” and “typically arises outside of science research settings routinely subjected to biosecurity and biosafety oversight and is typically undertaken by individuals unfamiliar with the history of biosafety guidelines.”¹⁷ It is that theoretical realm in which the concerns of this article lie.

EVALUATING RISK FROM AI IN THE CYBERBIOSECURITY SPACE

Artificial intelligence is an applied composite of basic sciences and applied disciplines, such as cryptography, that have been recognized as a set of tools that may have military and nonmilitary uses. As described

in their 2018 report, Hoadley and Lucas¹⁹ suggest that AI will present challenging applications for defense intelligence/surveillance/reconnaissance, logistics, cyberspace management, command and control, autonomous vehicles, lethal autonomous weapons systems, and the evolution of combat. The methods that might be used to create these superiorities range from improvements in data acquisition and data processing to improvements in modeling methods, such as the adaptation of machine learning for natural language or image processing, neural networking for machine vision, or AI-enabled software that can detect artificially intelligent viruses.

Whether applied to the problem of data acquisition and processing or modeling, AI stands to become a potentially hazardous tool at many points in the use of the technology. Yampolskiy suggests that there are eight scenarios that make up the “pathways to dangerous AI” (see Table 1, pp. 143–149, in Yampolskiy²⁰). He proposes that there are three types of external causes for dangerous AI: purposeful malicious acts, mistaken actions, and malice that arises from a “complete package from some unknown source.” He also proposes an internal cause of malice that may arise when a system takes on independent capacities. These causes may arise during either a pre- or a postdeployment phase.

As it is difficult to determine what about AI is specifically dual use (this is the case for science generally), the approach taken to AI risk should start from the view that there are multiple ways in which AI could be turned (in the short or long term) to malevolent uses if directed to data having implications that are harmful by definition

(for example, missile guidance) or could be combined to create novel harms. While most AI applications do not have the potential to take the “treacherous turn,” AI endowed with reinforcement learning or self-improvement components may learn adverse lessons from its experiences with data that could be directed to harmful ends and, consequently, pose a somewhat greater potential for harm. As described previously, the data related to some biological agents (for example, select agents and toxins) are “always already” a dual-use concern, with or without AI. What AI introduces into this cyberbiosecurity setting is the possibility that harmful uses could be identified more quickly or that lethality could be augmented by AI identifying sequences of genetic material, encoded in DNA or RNA, more quickly. Murch clarifies some of these concerns:

From protecting high-value intellectual property or sensitive personal health information, to ensuring that critical medical instrumentation and equipment is impervious to cyberattack so patients are not put at risk, to maintaining the integrity of genomic data being shared in “the cloud,” to protecting against the disruption or takeover of agricultural systems, to ensuring that advanced manufacturing produces what is expected with no unintended consequences, threats and risks exist now that need attention and could present and cause crises from which it would be difficult to overcome and recover.¹

As with other areas of research, measurement of the influence of AI in

TABLE 1. Classifying risk for AI applications.

Risk level	Data	Easily explained		Difficult or impossible to fully explain	
		Big data analytics	Supervised machine learning	Unsupervised machine learning	Advanced neural network methods
Low risk to national security	Health or behavioral data on humans or biological systems that could be exploited to create harms to health or welfare that do not include those from biological weapons	Personal health information linked to spending and employment data to create widespread targeted advertising of ineffective or dangerous preventive or curative products	Detection and classification algorithms trained, whether intentionally (bad actor) or unintentionally (faulty data labeling), to misidentify images in scans or X-rays, leading to false-negative or false-positive diagnoses	Clustering algorithms intentionally crafted to present larger or smaller estimates of symptom clusters or treatment efficacy	Polymorphic malware used to infiltrate a database to feed adversarial examples to neural networks built for classifying disease, leading to false-negative or false-positive diagnoses
High risk to national security	Health or behavioral data on humans or biological systems, including pathogens of known risk that could be exploited to create harms that include biological weapons	Linking personal health information, public health surveillance information, weather conditions, supply chain data, and demographics to identify areas of lower preparedness or higher susceptibility to lethal attack	Use of specialized algorithms to identify local conditions for augmentation of an agent's pathogenicity	Use of clustering algorithms to identify clusters of genes encoding for specific toxins or proteins in organisms with unknown pathogenicity	Adversarial examples fed into distributed neural network systems to train algorithms to misclassify engineered genetic sequences, thus confounding efforts to create effective medical countermeasures

biological research and development, including that with biosecurity implications, is a matter of risk assessment. The work reviewed in this article is offered as an argument to guide the development of tools for the evaluation of AI-enhanced cyberbiosecurity risk. Using a combination of Yampolskiy’s “pathways to dangerous AI” and perspective DURC that can relate to both materials and methods, we suggest that a materials-focused way of assessing biological data risks can be coupled with a methods-of-AI risk assessment.


Not all biological data pertain to agents with biosecurity implications. However, data from genomic studies, ancestry databases, personal

health-care information, gene-screening libraries, databases of biological agents’ sequenced genomes, and findings from mutational outcomes of CRISPR analyses could be brought together for myriad uses, both benign and malicious. Likewise, not all algorithms pose risks to individuals whose data are subject to them, and not all methods of machine learning could yield conclusions with adverse consequences for human or biotic systems’ safety. However, focusing on only “known” risks or “explainable” methods may leave open the possibility that unforeseen risks catch the scientific community off guard, as members of the DURC committees suggest is possible. In Table 1, we bring these together

as a way of identifying examples of potentially risky interactions.

As described by the Committee on DURC, “There is a growing tension between a scientific culture based on transparency and the need for secrecy to protect national security. . . . The ideal of a scientific culture based on principles of openness and transparency faces continuing challenges. One challenge relates to a concern that adversaries might take advantage of advances in science and technology for malicious purposes.”¹⁷ Open access to biological data is a significant reason that science has had the capability to

advance as rapidly as it has. Likewise, open access to AI code, packages, and implementation has led to the advancement of AI to its present state of near ubiquity.

While it is understood that these open access public databases and data repositories are vital to scientific operations, these should be paired with additional efforts to raise awareness of the increase of dual-use potential when these databases, codes, and associated publications, which offer detailed methods and materials for creating deadly agents, become publicly available. The risk assessment tools described previously are one element of the move toward a more aware and ethically informed science of AI-enabled cyberbiosecurity. 

REFERENCES

1. R. S. Murch, W. K. So, W. G. Buchholz, S. Raman, and J. Peccoud, "Cyberbiosecurity: An emerging new discipline to help safeguard the bioeconomy," *Front. Bioeng. Biotechnol.*, vol. 6, pp. 1–6, Apr. 2018. doi: 10.3389/fbioe.2018.00039.
2. S. E. Duncan et al., "Cyberbiosecurity: A new perspective on protecting U.S. food and agricultural system," *Front. Bioeng. Biotechnol.*, vol. 7, p. 63, pp. 1–7 Mar. 2019. doi: 10.3389/fbioe.2019.00063.
3. "Recommended data repositories," *Nature*, 2019. [Online]. Available: <https://www.nature.com/sdata/policies/repositories>
4. A. Baravalle, M. Sanchez Lopez, and S. W. Lee, "Mining the dark web: Drugs and fake IDs," in *Proc. 2016 IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, pp. 350–356. doi: 10.1109/ICDMW.2016.0056.
5. D. B. Resnik, "H5N1 avian flu research and the ethics of knowledge," *Hastings Center Rep.*, vol. 43, no. 2, pp. 22–33, 2013. doi: 10.1002/hast.143.
6. A. Paschke, D. Dimancesco, T. Vian, J. C. Kohler, and G. Forte, "Increasing transparency and accountability in national pharmaceutical systems," *Bull. World Health Org.*, vol. 96, no. 11, pp. 782–791, Nov. 1, 2018. doi: 10.2471/BLT.17.206516.
7. D. B. Resnik, "Research ethics," in *International Encyclopedia of Ethics*. Hoboken, NJ: Wiley, 2018, p. 3. doi: 10.1002/9781444367072.wbiee001.pub2
8. Panel on Scientific Responsibility, *Responsible Science: Ensuring the Integrity of the Research Process*. Washington, D.C.: National Academies Press, 1992, p. 48.
9. "Singapore statement on research integrity," World Conference on Research Integrity, Singapore, Sept. 22, 2010. [Online]. Available: <https://wcrif.org/guidance/singapore-statement>
10. "Code of ethics and professional conduct," Association for Computing Machinery, New York, p. 3. [Online]. Available: <https://www.acm.org/code-of-ethics>
11. L. Floridi and J. Cowls, "A unified framework of five principles for AI in society," *Harvard Data Sci. Rev.*, 2019, vol. 1, no. 1. doi: 10.1162/99608f92.8cd550d1. [Online]. Available: <https://hdsr.mitpress.mit.edu/pub/10jsh9d1>
12. P. Gandhi, "Explainable artificial intelligence," *KNuggets*, Jan. 2019.

ABOUT THE AUTHORS

SARA R. JORDAN is the policy counsel for artificial intelligence at the Future of Privacy Forum and affiliate faculty with the School of Public and International Affairs at Virginia Tech. Her research interests are in risk assessment and governance for data sharing, particularly for artificial intelligence applications. She is a Member of IEEE. Contact her at srjordan@vt.edu.

SAMANTHA L. FENN is a project coordinator for ATCC's Federal Select Agent Program in Manassas, Virginia. Her research interests integrate science and public policy to safeguard the bioeconomy by raising national security concerns regarding opportunities for the misuse of biological data for nefarious acts. Fenn received a master of public administration from Virginia Tech. Contact her at slfenn@vt.edu.

BENJAMIN B. SHANNON is a senior analyst with the U.S. Federal Bureau of Investigation. His research interests include homegrown violent extremists, specifically their ability to exploit human capital and intelligence capital to bridge physical acts of terrorism and cyber-based acts of espionage. Shannon received a master of public administration from Virginia Tech. Contact him at bshannonvt@gmail.com.

- [Online]. Available: <https://www.kdnuggets.com/2019/01/explainable-ai.html>
13. Z. Obermeyer and E. J. Ezekiel, "Predicting the future: Big data, machine learning, and clinical medicine," *New Engl. J. Med.*, vol. 375, no. 13, pp. 1216–1219, 2016. doi: 10.1056/NEJMp1606181.
 14. S. R. Jordan, S. L. Fenn, and B. B. Shannon. Feb. 8, 2019, Interview with Edward You, unpublished.
 15. "Safeguarding the bioeconomy: Applications and implications of emerging science," National Academy of Sciences, Washington, D.C., July 27–28, 2015. [Online]. Available: https://www.ehdc.org/sites/default/files/resources/files/Safeguarding%20the%20Bioeconomy_II_Recap%20Final%20090815.pdf
 16. Institute of Medicine, National Academy of Sciences, and National Academy of Engineering, *Scientific Communication and National Security*. Washington, D.C.: National Academies Press, 1982. [Online]. Available: <https://doi.org/10.17226/253>
 17. Committee on Dual Use Research of Concern. *Dual Use Research of Concern in the Life Sciences: Current Issues and Controversies*. Washington, D.C.: National Academies Press, 2017, pp. 60–63.
 18. Board on Life Sciences; Division on Earth and Life Studies; Committee on Science, Technology, and Law; Policy and Global Affairs; Board on Health Sciences Policy; National Research Council; and Institute of Medicine. "3: Gain of-function research: Background and alternatives," in *Potential Risks and Benefits of Gain-of-Function Research: Summary of a Workshop*. Washington, D.C.: National Academies Press, Apr. 13, 2015. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK285579/>
 19. D. S. Hoadley and N. J. Lucas, "Artificial intelligence and national security," Congressional Research Service, Washington, D.C., Rep. R45178. 2019. [Online]. Available: <https://fas.org/sgp/crs/natsec/R45178.pdf>
 20. R. V. Yampolskiy, "Taxonomy of pathways to dangerous artificial intelligence," in *Proc. 2nd Int. Workshop AI, Ethics and Society*, 2016, pp. 143–148. [Online]. Available: [arXiv:1511.03246v2](https://arxiv.org/abs/1511.03246v2)

IEEE Computer Society Has You Covered!

WORLD-CLASS CONFERENCES — 200+ globally recognized conferences.

DIGITAL LIBRARY — Over 780k articles covering world-class peer-reviewed content.

CALLS FOR PAPERS — Write and present your ground-breaking accomplishments.

EDUCATION — Strengthen your resume with the IEEE Computer Society Course Catalog.

ADVANCE YOUR CAREER — Search new positions in the IEEE Computer Society Jobs Board.

NETWORK — Make connections in local Region, Section, and Chapter activities.

Explore all of the member benefits at www.computer.org today!



Digital Object Identifier 10.1109/MC.2020.3020204

