The Weaponized Web
The National Security Implications of Data

Tech Policy Through the Lens of National Security
By Lindsay Gorman, Bret Schafer, Clara Tsao, and Dipayan Ghosh
**Executive Summary**

The permissive and sector-based data governance laws that have shaped technological innovation in the United States have constituted both a tremendous boon to U.S. innovation and a growing vulnerability in our national security. Authoritarian states, like Russia and especially China, have made control over data, both domestic and foreign, a centerpiece of their global strategies. The EU has approached data very differently, adopting the General Data Protection Regulation (GDPR) to establish rules governing firms that use the data of European citizens. The United States’ approach has so far been piecemeal, but the very openness that has been an engine of U.S. technological development is also increasingly weaponized by authoritarian powers who see advantage or opportunity in the control and abuse of Americans’ data.

To meet this threat, the U.S. government will need to adopt reforms to its regulatory approach that guarantee privacy and data integrity while also preserving the openness that facilitates experimentation and innovation. Important regulation has taken place on the state level, as state governments like California and Virginia have refused to wait for federal authorities to act. The following proposals breakdown the data governance challenges into five actionable steps that Congress and the administration can take at the federal level to improve U.S. national security, while protecting user data and providing a predictable and manageable regulatory regime for industry.

**Proposals**

1. **Require third-party data brokers to register with the FTC (or a newly established data protection authority), pay annual registration fees to fund enforcement, disclose ties to foreign governments and corporations, and establish limits on the types of data that can be sold to third parties without explicit user consent**

   Data brokers gather users’ information into datasets that have commercial value to their clients. But this data can also frequently be of value to foreign intelligence services, as well as malign non-state actors. Though firms are barred from selling to such groups directly, bad actors use straw-purchasers and intermediaries with impunity because the market remains unregulated. Requiring FTC or other registration would be a strong first step to protect the integrity of Americans’ data.

2. **Limit the acquisition and sale of biometric and genomic data (e.g., facial recognition images, fingerprints, DNA) to exclude certain covered foreign entities**

   Certain individually identifying and immutable biometric data, including facial structure/mapping, DNA, and fingerprints, are so fundamentally sensitive as to be worthy of special scrutiny when regulating the use, sale, and storage of data. Authoritarian states have placed a high premium on such data because of its value to the tools and tactics of authoritarian social control. The U.S. owes it to its citizens and residents to protect this data, and to require firms operating in the U.S. to do the same. This means preventing the sale or transfer of biometric data to foreign entities likely to abuse it for violations of rights and civil liberties.

3. **Require more of companies that amass citizen data by tying cybersecurity requirements to the amount and type of data collected, processed, or stored**

   Preserving an innovative environment while improving security can be aided by developing a tiered approach to data regulation that requires higher standards of firms dealing with larger quantities of data. Such an approach permits experimentation and development while incentivizing firms to build data security into their growth plans.

4. **Pass a small business cybersecurity tax credit**

   Such a tax credit would incentivize firms to invest in cybersecurity from the beginning, making them more likely to make it a priority as they grow, and spur growth in a critically important cybersecurity market.
5. Require companies to notify a federal authority, such as CISA, of a breach, including the type of incident, soon after the company became aware of the intrusion

Current law, at the state level, only requires firms that have suffered a data breach to alert authorities if the breach compromised PII. This means that many incidents, some of national security interest to the United States, go unexamined by investigators, researchers, and regulators. To establish a clear threat picture of authoritarian cyber actors and incidents, notification for data breaches beyond those containing PII is needed.
Open societies have cultivated rapid technological advancement and market innovation—which have vastly outpaced democratic governance. Authoritarian powers have seized on the underlying opportunity to exploit the open standards of the democratically regulated digital information environment and undermine democratic values and institutions while shoring up their own regimes. This poses a novel challenge for democracies, which must adapt to compete in this conflict over the data, architecture, and governance framework of the information space without compromising their principles. Effectively competing with autocrats in this environment—and ensuring a democratic future for the online information space—will require policymakers to analyze technology and internet policy through the lens of national security.

This paper analyzes policy options to mitigate national security concerns related to the collection, retention, and processing of data. The report is narrowly scoped in order to provide sufficient space to debate the relative merits of regulatory proposals. Although the authors make specific recommendations, the purpose of this report is to foster deeper debate about potential policy options; as such, we present arguments for and against each regulatory option, with an emphasis on highlighting potential negative externalities. While intended for a global audience, the recommendations are decidedly U.S.-centric, due simply to the authors’ deeper understanding of the U.S. regulatory landscape.

It is worth noting that this paper does not address the full spectrum of consumer-centric and privacy arguments for regulating the data industry. We also recognize that government regulation is not the only—and often not the best—tool to solve many issues in the digital domain. Our goal, therefore, is not to solve the myriad issues with how we produce, distribute, and consume information—it is to assist regulators and concerned stakeholders in thinking through legislative options to mitigate the national security concerns associated with malign foreign activity, interference, and alternative modes of governance in the technology domain. This focus means that our recommendations are more narrowly scoped to the national security challenge than those advanced by others—including the Digital Innovation and Democracy Initiative also housed at the German Marshall Fund.

Finally, it is our hope that this report will move the tech policy conversation beyond empty platitudes, generalities, and well-intentioned but ultimately impractical proposals. Years after these problems first surfaced, it is our shared belief that it is time to stop admiring the problem and focus instead on concrete solutions.
A Fragmented U.S. Approach

The United States is one of only a few developed nations without a comprehensive law regulating the collection, use, and storage of consumer data, opting instead for a sector-based approach in which federal privacy laws apply to specific contexts (such as health care, financial services, education, and children’s information). This approach has enabled the explosive growth of digital services companies—from tech giants to unknown data brokers—that have profited from the capture, analysis, and commoditization of personally-identifiable information (PII). It has also led to widespread data protection failures, abuses of privacy, and concerns over the consequences of predictive algorithms and surveillance capitalism for democracy and human rights.

To date, however, Congress has failed to address these vulnerabilities through federal data security or privacy legislation. There is no national data and privacy protection law in the United States that standardizes user data, collection, retention, and sharing requirements, although several existing statutes address elements of data access and privacy. The U.S. Privacy Act of 1974 gives American citizens the right to access any data held by government agencies and the right to copy that data. The Health Insurance Portability and Accountability Act (HIPAA) of 1996 provides data confidentiality requirements to allow for privacy and permissions around health data held by traditional medical and insurance providers. And the Children’s Online Privacy Protection Act (COPPA) of 2000 regulates personal information collected from children under 13. However, the lack of a modern, omnibus user data privacy and protection law (or authority to enforce violations) has left the onus on states to protect Americans’ information from cyber-enabled threats.

In 2002, for example, California passed the first law requiring notification for cyber events that resulted in the compromise of certain kinds of personal information. By 2018, all major U.S. jurisdictions (the fifty states, District of Columbia, and Puerto Rico) had data breach laws that covered information such as Social Security numbers and payment card information. After the implementation of the General Data Protection Regulation (GDPR) in the European Union in 2018—a law that imposes obligations on companies and organizations anywhere that target or collect data related to people in the EU, with the goal of giving EU citizens more control of their data—several U.S. states filled the federal void to pass their own comprehensive data privacy laws. The California Consumer Privacy Act (CCPA), passed in 2018, afforded California residents some, though not all, of the protections provided by GDPR. Most notably, it adopted an expansive definition of protected personal information—beyond traditional, narrow types of data such as social security numbers or payment information—as well as the ability to opt out of the sale of personal data and the right to sue in the event of a data breach. Nevada and Virginia followed suit with their own data protection laws, and more than a dozen other states are currently considering some type of data privacy legislation that incorporates aspects of the European or California laws.

There have also been limited attempts by some U.S. states to regulate specific elements of the big data business model. Several states have enacted legislation governing the use of biometric information. The most notable impact to date has come from the Illinois Biometric Information Privacy Act. Enacted in 2008, the law allows Illinoisans to seek damages from companies that improperly collect or store their biometric information. New York and Washington also have biometric information protection laws. And in 2019, Vermont enacted the nation’s first data broker registration law, requiring third party data brokers to register with and provide annual disclosures to the Secretary of State. California followed suit in 2020 with a data broker registration law, and more than a dozen other states are currently considering some type of data privacy legislation.

Actions at the state level show both the need and the appetite to regulate the “Big Data” industry. But leaving regulation to individual states has created a confusing regulatory landscape that is difficult for companies, let alone average citizens, to navigate. And while states have an important role to play in consumer protection, they are ill-equipped and ill-resourced to the challenge of defending Americans and American interests against the national security risks posed by foreign threat actors.
Momentum for Federal Legislation

If there is a silver lining, however, it is that the threat of an even more fragmented regulatory environment has provided momentum, both at the industry and government level, to create a federal law to ease the complexity of data compliance. Much of this regulatory energy has been understandably contextualized in terms of the established role of government regulators in protecting consumers from harms—be it unsafe automobiles, toxic toys, or predatory lenders. While there is ample evidence that data misuse has created similar real-world harms for consumers, this paper makes the argument that there is also a national security prerogative for enacting stronger data protection regulations.23

Lax data protection standards create numerous national security vulnerabilities. The acquisition—legal or otherwise—of big data or sensitive personal data creates the potential for foreign governments or intelligence services to surveil U.S. citizens, track intelligence assets or military personnel, develop targeted biothreats, perfect AI systems, or compromise influential figures, among other threats.

As China's surveillance reach extends beyond its borders, the ability of individuals to operate freely and openly in democratic countries like the United States is increasingly challenged. The popular all-purpose chat app, WeChat, for example, has been shown to surveil users’ communications even outside of China.24 25 Widely popular with the global Chinese diaspora community, the app also censors sensitive political topics in conversations between individuals inside and outside of China.26 In addition to shaping the information realities of its users, apps like WeChat can have a chilling effect on free speech in the United States by providing the ability to target individuals based on their communications. While the risk may be minimal for those not planning to travel to authoritarian countries, for U.S. persons with travel plans to China or who have relatives in China, the potential cost of their speech or political activity in the United States to themselves or loved ones risks creating an environment of self-censorship. This risk is by no means hypothetical. In 2019, University of Minnesota student Luo Daiqing returned home to Wuhan, China after finishing his spring semester and was detained for months and then imprisoned by Chinese authorities. The cause? Tweets “denigrating a national leader's image” that Luo had posted months earlier while studying in the United States.27 If the United States seeks to remain a safe haven for free speech and freedom from authoritarian political influence, it has to guard against the risks that authoritarian data collection poses for those on U.S. soil.28 Moreover, as the United States seeks to build a coalition of its allies around a democratic vision of emerging, data-driven technologies, it will benefit from a robust, big picture data strategy that moves beyond the consumer lens and also considers the global regulatory landscape.
Policy Proposals

In this section, we outline a series of five legislative policy proposals to address national security vulnerabilities related to the collection, use, security, and transfer of personal data. For each, we provide background on the policy discussion, review arguments against and in favor of the proposal, and make a final recommendation to Congress.

1. Require third-party data brokers to register with the FTC (or a newly established data protection authority), pay annual registration fees to fund enforcement, disclose ties to foreign governments and corporations, and establish limits on the types of data that can be sold to third parties without explicit user consent

Background

Third-party data is consumer information that is purchased or sold by brokers that do not have a direct relationship with the consumer. Often, these brokers amass and aggregate data from multiple sources, allowing for the creation of complex, data-driven profiles that can be used for consequential purposes ranging from determining one's suitability for a job to shaping one's political opinions. Unlike websites and social media platforms where users consent—at least tacitly—to the collection of their data as a condition of their use of a service, data brokers collect and sell data from individuals who do not use or benefit from their services.

From a national security perspective, there are multiple risks associated with the largely unregulated practice of data brokers amassing and selling large and/or sensitive datasets. Perhaps the biggest threat is the risk of state or non-state hacks of internal servers that can compromise troves of sensitive data. The massive 2017 hack of Equifax, for example, was determined to be a Chinese military intelligence operation to steal trade secrets and personal data on more than 145 million Americans.29 While the risk of cyber intrusions exists across the data industry, third-party brokers are prime targets due to the volume of data they collect and their cybersecurity protocols, which are at times substandard.30

But foreign intelligence services can also gain access to sensitive data through entirely legal means. While it is illegal for brokers to sell data directly to sanctioned entities, which may include foreign intelligence agencies or officers, it is entirely legal to sell data to commercial entities that may have direct, indirect, or undisclosed ties to foreign governments. As regulators increasingly scrutinize Chinese and Russian-owned apps, which can have links to their respective national security and intelligence services, data brokers offer authoritarians a potential backdoor to American data.

Yet the push to limit Chinese and Russian-owned firms from directly accessing American data has not resulted in commensurate efforts to restrict American companies from wittingly or unwittingly providing that data through an opaque web of third-party data transfers. Globally, regulations like the EU’s GDPR31 and Singapore’s Personal Data Protection Commission (PDPC) limit, and in some cases prohibit, the collection of certain types of data by third-party brokers. But in the United States, only a small number of states have considered or enacted legislation to regulate data brokers, and those laws typically focus on transparency—requiring registration and in some cases the filing of reports—rather than imposing restrictions on the sale or transfer of personal information. Vermont’s data broker law also requires data brokers to maintain a comprehensive information security program to protect personal information and to register annually with the state’s Secretary of State.

But efforts to enact similar legislation at the federal level have largely stalled, despite bipartisan support. Most recently, the Data Broker List Act of 2019, introduced by Senators Gary Peters (D-MI) and Martha McSally (R-AZ), died in Congress without receiving a vote. That bill would have required data brokers to join a national data broker registry, overseen by the Federal Trade Commission, and maintain a comprehensive information security program.32 33 Regulation would have also required third parties to process data in a matter that is consistent with
users’ expectations of its intended purpose.  

Despite initial opposition, many within the data brokerage industry have now signaled their support for regulation. In a New York Times op-ed, the Chief Data Ethics Officer of Acxiom, an American database marketing company, argued that such a bill would provide “transparency, uniformity, and certainty” across the industry, without negatively impacting innovation and competition.  

Arguments Against  

Positive use cases of third-party data sales. Third-party data is an integral part of the digital ecosystem and vital to industries that rely on data as a core component of their businesses, including banks that use third-party data to detect financial fraud. The sharing of data can also create positive externalities in the fields of health, safety, and transportation (to name but a few) that in some cases grow as more parties share data.  

Consumers don’t care. Many leading data brokers already offer consumers the ability to opt-out of data collection, yet few people take advantage of this opportunity.  

Low state-level compliance and an unclear foreign purchasing threat. While requiring data brokers to join a federal registry would add a modicum of transparency to the industry, it would do little to solve the issues of data breaches or intentionally deceptive business practices absent meaningful enforcement. Vermont’s law, for example, has produced few cases and compliance has been low. Additionally, there have been few documented cases of data brokers selling sensitive information to foreign governments, suggesting a better approach is to focus on improved cybersecurity to prevent hostile cyber intrusions rather than adding an additional layer of bureaucracy.  

The U.S. government can also buy this data. Legislation in this space should also address whether U.S. government entities should have access to sensitive, aggregated third-party data without receiving explicit user consent. This tension between evolving conceptions of privacy and the public square and the government’s ability to access aggregated data for security needs should be resolved if the United States is to take a leading role in data privacy standards.  

Arguments in Favor  

Provides a window into who has Americans’ data and how it is used. A federal data broker registry would allow consumers and government watchdogs to better understand how American data is being bought, sold, analyzed, and aggregated. It is a necessary starting point to identify the companies that harvest Americans’ data—often without their direct knowledge or consent—in order to hold them to account for unethical or unsound business practices that not only jeopardize individuals’ privacy but also U.S. national security.  

We do not know if data brokers sell to authoritarian-connected entities. Mandating disclosures of ties to foreign corporations and governments is a sound national security practice that would highlight potentially problematic linkages and is in keeping with disclosure requirements imposed on other sensitive industries. Without those disclosures, it is exceedingly difficult for the private sector and government to understand the scope of the risks associated with selling data to certain entities.  

Deterrence measure against non-compliance. It is necessary to create disincentives (either through fines for non-compliance or a private right of action) to deter companies from selling data without proper notice and consent, particularly to authoritarian-connected entities, or from failing to safeguard sensitive data.  

Guards against data asset fire sales to authoritarian-connected entities. A registry provides an ability to start to track the ‘Big Data’ sector, and also gives a partial answer to the question of what happens to the data when a data broker business fails or goes bankrupt. Fire sales of failed data companies represent a vulnerability not covered by the Committee on Foreign Investments in the United States (CFIUS) process because they are not investments. If a foreign actor started buying up the data assets of failed companies, the U.S. government would have
no way of knowing or assessing the national security value of that data. As the data industry grows, a registry can provide indicators of trends and greater foresight if these companies start failing.

**Final Recommendation**

The Vermont and California data broker laws provide a roadmap for federal legislation and would help to identify companies whose activities relating to the sale of personal information could pose a national security threat—either through relationships with authoritarian governments or connected entities, misuse of consumer data, or the absence of comprehensive information security programs. Given the need for and the cost of enforcing compliance, we recommend that legislators require that data brokers pay a nominal annual data protection fee, similar to the fees mandated by the United Kingdom’s Data Protection Act. These fees would allow the relevant authority (whether the FTC or a newly created data protection agency) to commit the resources necessary to enforce compliance. Additionally, providing a private right of action for negligent, reckless, or intentional violations would also incentivize compliance. However, the potential for class action lawsuits to flood the federal court system is a legitimate concern. It is therefore our recommendation that the right of action be fairly restrictive, following the lead of California’s Consumer Privacy Act. Finally, the issue of the conditions under which the U.S. government should be able to purchase aggregated third-party data deemed too sensitive for foreign governments is of critical importance, but is a question for constitutional scholars and ultimately outside the scope of this paper.

2. Limit the acquisition and sale of biometric and genomic data (e.g., facial recognition images, fingerprints, DNA) to exclude certain covered foreign entities

**Background**

Authoritarian actors, China chief among them, have seized on the power of data as an input to artificial intelligence systems, a driver of economic opportunity, and a tool for information control. The Chinese Communist Party aims to collect and amass the world’s data, often without initial regard to how it will be used. The United States and its democratic peers have only just begun to contemplate what this data vacuuming effort means for the privacy of its citizens and of U.S. persons, particularly when it comes to sensitive or immutable personal data, such as biometric information. Until recently, collecting, processing, and analyzing biometric data at scale was not possible, and so specific laws governing data protection for this type of information were less needed.

In February of 2021, the National Counterintelligence and Security Center published a brief on “China’s collection of genomic and other healthcare data from America,” describing implications for privacy and U.S. national security. They included the risks we identify here of “vast opportunities to precisely target individuals in foreign governments, private industries, or other sectors for potential surveillance, manipulation, or extortion.” In addition, PRC companies use biometric information to train surveillance systems that target individuals by race, ethnicity, and other immutable qualities, which in turn fuels repression and undermines universal values around the globe.

Looking to the future of the information environment, the last two years have seen rapid technological progress on the creation of “deepfakes”—videos, images, or voice recordings manipulated with artificial intelligence (and specifically deep learning) to put words in someone’s mouth or create audio or visual realities that never happened. The mass, unregulated collection of biometric information opens new possibilities for fabrication of kompromat based on biological data, with added complications for authentication and provenance.

CFIUS provides one tool to block a foreign company from obtaining potentially sensitive U.S. citizen data, and the Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA) expanded its scope and powers. Over the past few years, CFIUS has begun to block transactions involving Chinese acquisitions of companies that have access to potentially sensitive U.S. persons’ data, including LGBTQ social networking app Grindr, and a hotel property management system StayNTouch. Grindr, for example, asks for sensitive information on HIV status and testing history. While this case-by-case approach has served to limit some transactions, it has created
confusion for businesses about what criteria would trigger national security concerns.

For consumers, the situation is even less clear. In the summer of 2019, for example, FaceApp, a “face-ageing” app of Russian provenance went viral in the United States, attractive for its ability to simulate the ageing process on photos of user-supplied faces. Its authoritarian provenance raised concerns, including from U.S. Senate Majority Leader Chuck Schumer about how the data of American faces could be used by the Russian government. The ensuing FBI investigation assessed in December of 2019 that the fact it was built in Russia made the app a “potential counterintelligence threat,” but provided little guidance to the consumer on how to manage or understand these risks.45

A more effective approach to regulating these transactions would be one that clarifies the types or amounts of data whose sale to covered foreign entities in authoritarian countries is restricted, or whose provision to these entities raises concern and warnings. Such an approach would place guardrails against the reality that sensitive data on U.S. persons continues to end up in the hands of authoritarian actors that could use it for blackmail, influence, or manipulation.

A data-sharing framework with more clarity could be internationalized to democratic partners looking to resist authoritarian influence. In some areas, the building blocks for such an approach may be found outside the United States. The EU’s GDPR, for example, affords biometric data a special status even among sensitive data, requiring that data controllers conduct privacy impact assessments for many ways of processing biometric data.46

Arguments Against

It’s what China does. Democracies have traditionally championed the open data model, in sharp contrast to emerging authoritarian internet norms that are characterized by information control, data sovereignty, and data localization. At the core of democratic governance is a belief that information should flow freely, and that restrictions on access to it are the province of despots. China, for example, has long held data localization requirements that specify that personal data and national security-sensitive data developed in China stay in China. Beijing is also developing a system for classifying data based on its sensitivity. There is a risk that the United States goes too far in emulating this approach and sacrifices both principles and innovative potential in an attempt to control risks that can be speculative or ill-defined. In short, the risk calculus of this move may be off.

Sensitivity by aggregation. Classification of data as sensitive or non-sensitive is the wrong metric in a world where aggregated data is used to draw inferences with accuracy. A piece of data on its own may not be inherently sensitive, but the inferences about individuals or about a collection of individuals that are drawn from collections of non-sensitive data can be. For example, an instantaneous piece of location data yields little insight into an individual. But 24 hours of location tracking from a cell phone is nearly a precise personal identifier.47 This sensitivity by aggregation is well understood in the intelligence community, where classifications can be upgraded due to the aggregation of information. Defining individual pieces of data as sensitive or non-sensitive, and building frameworks around them, may be of limited utility.

Scoping and defining biometric data. There are difficulties in defining what counts as biometric data. Voice prints, for example, could be used to identify individuals, but are becoming increasingly popular through platforms like Clubhouse and Twitter. Advances in machine learning have created an interest in analyzing speech patterns for improved voice recognition. As biometric data becomes increasingly part of social discourse, regulating it may be more challenging.

Balkanization of the internet. Finally, some argue that the adoption of regulations on data and technology in democracies would lead to a further balkanization of the Internet. The argument is that a free and open internet is the best guarantor of universal digital rights, and that a splintering of the internet and technology ecosystems along geopolitical lines risks further economic cleavages that ultimately entrench authoritarian norms.
Arguments in Favor

Regulatory clarity and innovation potential. Through recent CFIUS decisions that have required companies involved in processing U.S. persons data to divest from Chinese ownership, the United States has already signaled a willingness and a desire to stop some forms of mass data collection in the United States by companies linked to authoritarian countries. This defensive approach provides an after-the-fact stopgap for data collection platforms that already pose a national security risk to the United States. It does not, however, give guidance to U.S. firms in data-driven industries. It also only addresses investment and ownership risks—not the transfer of data itself. A more proactive approach that anticipates risks can create a clearer innovation environment for biometric, genomic, and health care data industries. Limiting exports of sensitive data to certain nations could allow for domestic companies to use aggregated data in a way that is useful to business models that take full advantage of the data life cycle.

User guidance. In too many areas, cybersecurity and data security risks are passed on to consumers, who have little means of judging where their data is going and which information platforms, apps, or services will guard it. Clear limits would remove this responsibility and allow consumers to access services without concern that their DNA or biometric information may end up in the hands of an authoritarian regime. Conversely, while Americans are generally sensitive about personally identifiable information, like social security numbers and medical records, they have little recognition of the value of their DNA or other biometric information. Laws that privilege biological and biometric information protection and privacy could also play a role in raising public awareness and education on the value of this data.

Resisting long-term healthcare dependence. China in particular is seeking to amass genomic and healthcare data on citizens of the United States and countries around the globe. Its aim is to become the global leader in future healthcare industries of personalized medicine, vaccine development, and individually-targeted therapies. As Bill Evanina and Edward You have argued, success in future medicine and healthcare industries will depend in part on the ability to collect, store, and process the genomic data of individuals. Outsourcing U.S. biometric and genomic data—and eventually the healthcare services they spawn—to China presents a “long-term existential cost to our nation.”

Internet balkanization is underway anyway. Regardless of U.S. policy actions around data, the internet is already increasingly split because of the closed information systems in authoritarian countries. Multinational technology companies that harness user data already face difficult choices when seeking to operate in authoritarian countries where the rule of law is weak, and the state may demand censorship and access to data.

Defining biometric data is doable. There are precedents for defining biometric data, including the Illinois Biometric Information Privacy Act and the EU’s GDPR. While these definitions may need to be updated with an eye towards present and future uses of such data, the challenge is not insurmountable. Any legislative framework could either specify principles against which future developments could be weighed or provide for regular reconsideration of the definition over a period of 5 or 10 years.

Final Recommendation

The United States should establish a comprehensive bio data strategy that considers the full lifecycle of biological data, including the infrastructure to store, transfer, and analyze it, as well as protections for government, industry, and consumers. Part of this effort should include a framework that clearly defines risks, and limits the ability for covered entities to access biometric, genomic, or healthcare data on U.S. citizens. Covered entities are those based in or tied to authoritarian countries without credible independence from the government or a strong rule of law system. In addition to limitations on the transfer of data itself, a consideration of the harms should be part of this risk calculus. Additionally, Congress should consider amending and expanding HIPAA to extend the protections it affords health data collected by doctors, hospitals, insurance companies, and traditional medical providers to include consumer service companies that collect DNA or healthcare information. This action could provide a baseline level of protection of individuals’ sensitive health data in an economy where that data no longer stays with traditional medical care providers. The details of what such a proposed HIPAA expansion could
look like fall outside the immediate scope of this paper; given their economic and regulatory complexity in costs, oversight, and responsibilities, they should be addressed in separate work.

3. Require more of companies that amass citizen data by tying cybersecurity requirements to the amount and type of data collected, processed, or stored

Background

In December 2020, Congress passed the Internet of Things Cybersecurity Improvement Act of 2020, which codifies standards and guidelines for the federal government for the management of connected devices.49 While there is certainly a need for elevated data security protections for government officials, the need for increased cybersecurity standards extends to all companies that amass personal or sensitive data. Leaving cybersecurity to the discretion of individual companies creates the potential that some companies may choose to forego costly or labor-intensive security measures. In contrast to other business decisions best left to the market, security lapses from data collection companies potentially expose trade secrets or compromising information about individuals. This can be, and has been, used for state-sponsored corporate espionage and kompromat.

Moreover, the economic and geopolitical value of data may change over time in relation to how it is aggregated and how it is used. A state-of-the-art machine learning system, trained on robust and wide-ranging datasets, can become valuable intellectual property. But right now, the United States is not protecting a crucial building block of that intellectual property—its data—from authoritarian governments. In addition to corporate information, in the age of artificial intelligence, personal and biological data retains value beyond its use for identity theft.

Arguments Against

Costs to small businesses. The initial costs of such measures may be steep for small businesses and unfairly advantage larger firms with plug and play compliance structures.

Unclear responsibilities. Across the IT architecture chain, the scope of responsibility is unclear. Small- and medium-sized enterprises, for example, usually do not host their own data. In such cases, it is unclear whether the liability lies with the hosting provider or the entity that “controls” the data. There are also dynamics of power discrepancies between large data-storing companies and the usually smaller firms that use the data. Such dynamics could complicate the ability to develop fair regulation regarding cybersecurity burdens.

Unclear liabilities. Because data’s value changes over time and is not linked solely to identify theft, assessing penalties and civil liabilities for the failure to protect data is inherently challenging.

Existing requirements. Federal agencies like the FCC and SEC already establish certain cybersecurity requirements. In particular, the SEC states that publicly traded firms are supposed to report on cybersecurity risk. It also requires carriers that maintain information from subscribers to provide adequate security in terms of storage and retrieval. In 2018, the SEC issued guidance that publicly traded companies should periodically disclose “material cybersecurity risks and incidents in a timely fashion” to investors, weighing factors such as financial risk to investors and the importance of any compromised information.50

A roadmap for adversaries. Public disclosures of cybersecurity compliance would outline a roadmap for adversaries to target entities that are high-value and poorly defended.

Arguments in Favor

Corporate incentives are misaligned with long-term national security objectives. Short-term corporate and financial interests are misaligned with the robust cybersecurity protections necessary for safeguarding data from the standpoint of national competitiveness.51 Passing the cybersecurity risk onto the consumer may be acceptable as a business strategy when investors are focused on short-term gains. As a national security matter, it undermines long-term competitiveness. Like federally mandated safety requirements for nearly every major industry,
cybersecurity requirements for companies dealing with sensitive personal data should be required and enforced by federal authorities.

**Existing requirements and enforcement are minimal.** Current requirements, including those from the FCC and SEC, are minimal, loosely defined, and infrequently applied. A March 2021 report commissioned by SecurityScorecard, for example, found that publicly-traded companies fall far short even of the SEC’s 2018 guidance.\(^52\) Boiler-plate legalese such as “[c]yber-attacks could have a disruptive effect on our business” takes the place of rigorous risk analysis. And the study found that only 17 percent of Fortune 100 companies disclosed management-level cyber-related issues to their boards or relevant board committees at a “frequency of at least annually or quarterly.”\(^53\)

**Precedents already exist.** There are precedents for this approach at the state and international levels. In California, CCPA obligations are tied to company size, with two measures tied to the amount of data a company collects or sells. Many state data breach laws also impose cybersecurity obligations. The European Union Cybersecurity Act institutes a cybersecurity certification which can enforce certain standards, such as a frequently updated risk assessment.\(^54\) Requiring companies to report risk assessments, security protocols, and other relevant information to a centralized agency can increase compliance and data safety. There is significant agreement on the tools necessary for strong cybersecurity, but little means of centralizing them and requiring companies to act.

**Final Recommendation**

Congress should create a cybersecurity reporting structure, analogous to the SEC 10-K filings for publicly traded companies, for companies in the United States that compile, process, or store certain quantities of data. A data regulator should issue guidance for cybersecurity compliance based on firms’ data sizes (how much personal data they collect on users), develop a minimum standard of compliance, and launch a rigorous certification and recertification process. Congress should also explore establishing liability for breached companies that do not meet minimum standards.

**4. Pass a small business tax credit**

**Background**

As detailed in proposal 3, companies that handle data that could pose a national security risk if compromised should be required to implement certain cybersecurity standards. However, initial investment and ongoing maintenance costs to purchase cybersecurity technology and services could prove cost-prohibitive for small- and medium-sized businesses. Simply exempting small businesses, though, would be a flawed approach—sensitive data remains sensitive regardless of the size of the business holding that data.

Exempting small businesses from cyber security standards would also invite further malfeasance from malign cyber actors. Despite extensive coverage of major data breaches at large corporations (Equifax, Marriott, Facebook, etc.), more than half of small businesses reported a cyberattack in 2019,\(^55\) yet nearly 6 in 10 reported having no cybersecurity plan.\(^56\) This creates an obvious vulnerability, as small corporations that work with sensitive data or on sensitive topics often do not have the resources to defend against hostile state attacks. This is particularly true of nonprofits, nongovernmental organizations, and think tanks that work on human rights, foreign policy, national security, and other issue areas of interest to state-backed hackers. In 2018, for instance, the German Marshall Fund was one of several think tanks targeted by hackers linked to the Russian government.\(^57\) And in December 2020, the Cybersecurity and Infrastructure Security Agency (CISA) issued an alert warning that U.S. think tanks were being targeted by advanced persistent threat actors.\(^58\) But the threat is not limited to organizations involved in policy or national security discussions. Start-ups that hold sensitive personal or corporate data are also likely targets.

A possible solution that would encourage small businesses and nonprofits to invest in cybersecurity measures without putting them at a competitive disadvantage is to offer qualified companies a tax credit. The State of Maryland, for example, offers companies a tax credit of up to 50 percent of the net purchase price of cybersecurity
ty technologies and services (though service providers must also be a qualified Maryland company). A similar credit at the federal level would incentivize companies to invest in and maintain information security programs, and could be contingent on investment in technology, services, or software from American or other approved vendors. This would help to ensure that companies are not only investing in cybersecurity, but doing so through service providers that do not have potential connections to the very actors these measures are meant to defend against.

**Arguments Against**

**Potential for waste and artificially inflated costs.** With any tax credit, there is a concern about who would be eligible, how the credit would be funded, and the potential for waste and abuse. In addition, there is the potential to raise the pricing floor for eligible goods and services for as long as the incentive is in place, making it potentially more expensive for all companies to purchase cyber security technologies.

**Consumer education, reputational advantage, and market forces should drive compliance.** Market forces should be the driving force behind compliance. Better consumer education and mandated disclosures of companies’ cybersecurity practices and breaches would help to drive compliance without government intervention. In the long-run, businesses that invest in and maintain strong cybersecurity programs will be at a competitive advantage over those that do not, and the reputational risks and the costs of mitigating data breaches should be enough to incentivize strong information security programs.

**Lack of focus on the weakest link: humans.** The weakest links in companies’ information security chains are often their employees. A government program that promotes investment in information technology without the requisite training of staff might be ineffective and, ultimately, a waste of taxpayer money.

**Arguments in Favor**

**Insufficiency of the market to prioritize security.** In reality, market forces are unlikely to prioritize national security concerns. On the contrary, individual consumers often behave in ways that do not prioritize security over other factors (cost, features, convenience, etc.), and businesses only sometimes emphasize cybersecurity in their business-to-business transactions. Moreover, a company’s cybersecurity commitment is not always apparent, nor is there always a direct correlation between preparedness and subsequent cyber incidents that would inform any reputational calculus.

**Non-exclusionary of awareness training.** While humans are indeed the weakest link in cybersecurity, a tax credit could include broader support for vendor-provided cybersecurity awareness training or no-cost access to government-developed modules.

**Levels the playing field while raising the bar.** If cybersecurity standards are mandated at the federal level, a tax credit can level the playing field between the tech giants that can afford to meet those standards, and those for whom such an investment would be cost prohibitive. While private companies that amass sensitive data certainly have an individual responsibility to protect that data, the government can and should help mitigate expenses that come with adhering to government-imposed standards.

**Encourages trusted supplier purchases.** By incentivizing investments in American (or allied) cybersecurity technology, the government can encourage businesses to purchase equipment and services from trusted suppliers.

**Final Recommendation**

To ensure that small- and medium-sized businesses can purchase and maintain the technology required to safeguard sensitive data and remain competitive, Congress should offer a tax credit to qualified companies. This credit should be contingent on companies meeting or exceeding federal standards and investing in technology or services from American (or allied) companies. The benefits to both the American economy and American national security should offset concerns over costs, but the credit could also be funded (at least partially) by fines.
levied by the FTC against data brokers and other non-compliant tech companies.

5. Require companies to notify a federal authority, such as CISA, of a breach, including the type of incident, soon after the company becomes aware of the intrusion

Background

Currently, breach notification requirements exist in various forms at the state level centered on reporting breaches of personal data to authorities and/or consumers. This patchwork quilt of requirements creates disparities in terms of what constitutes a breach, who must be notified, what information that notification requires, and the timelines by which those notifications must occur. More importantly, these laws are aimed solely at consumer protection—not at forming a picture of foreign adversary intrusions into U.S. companies and datasets. Moreover, a focus only on personal information—while important—leaves the U.S. government in the dark on a range of national-security relevant cyber-attacks, including the scale and scope of intellectual property theft and business data exfiltration that can harm the competitiveness of U.S. firms in critical industries. Assessing a foreign government’s ability to hoover up data—personal, corporate, and otherwise—from across U.S. entities is critical to mounting defenses and competing in the information age. A federal breach notification law focused on understanding foreign adversary tactics, techniques, and procedures (TTPs), capabilities, and motivations when exfiltrating data could help the United States mitigate the national security consequences of data breaches.

Arguments Against

Legislation already exists. Breach notification legislation already exists widely at the state level. Not only would a federal law be duplicative in some cases, but it would create uncertainties over whether federal law would preempt state laws or other laws like HIPAA with their own reporting requirements.

Compliance requirements stifle competition. Additional reporting requirements to federal authorities inherently disadvantage small businesses without investigative and legal compliance structures already in place.

Pushback on government intrusion into private systems. To the extent that breach reports detail the workings of company systems or involve investigations with the U.S. government, some would argue that such actions constitute government overreach into private business.

Arguments in Favor

A more complete picture of nation-state data activity. The lack of notification requirements at the federal level makes it difficult for intelligence agencies to understand and assess foreign adversary TTPs. Without a holistic view of the threat landscape, relevant authorities are hampered in their ability to understand and respond to digital threats from nation-state actors.

Stemming the IP leakage to China. NPR has reported that the United States loses upwards of $57 billion from technology theft from China each year—a problem that is compounded by an unwillingness among companies to come forward and report cyber incidents. Corporate incentives not to alarm shareholders or to preserve business opportunities in China have led U.S. companies to refrain from reporting breaches. As such, federal prosecutors cannot hold perpetrators accountable, nor can the U.S. government understand where its advantages are eroding as a result of cyber espionage. Requiring notification to a federal authority of such incidents could help stem critical IP leakage by building an understanding of the threat and providing companies with the tools to protect their innovations.60

Current law can’t assess foreign actors. Existing breach laws are consumer laws. They do not require breached entities to explain TTPs, and states are not equipped to paint a full picture of foreign adversary cyber operations through breach notifications. This is the job of the federal government.
Final Recommendation

Congress should pass federal breach notification legislation that goes beyond the consumer protection purpose of public notification of personal information breaches. Because such a provision would represent an entirely new legislative purpose for breach notifications, Congress should consider the scope of such a law, as well as new types of information required in notifications, and new means of assisting attacked entities. In particular, notifications should be required to contain actionable information for CISA or a federal authority.
The authors would like to thank Susan Aaronson, Cameron Kerry, Ted Dean, April Doss, Harold Feld, Ayden Federline, Andrew Grotto, Ayane Miller, Adam Segal, and Edward You who participated in a roundtable that helped shape the arguments in this paper. Special thanks to April Doss, Maurice Turner, and Dave Salvo for providing crucial feedback. Nathan Kohlenberg, Joe Bodnar, and Kayla Goodson provided vital assistance with final production of the paper. Our greatest thanks and appreciation go to Ishmael Abuabara and Jenny Gurev who worked on all aspects of this project and were essential in bringing it to a successful conclusion.

Acknowledgments do not constitute an endorsement of the views and opinions expressed in this report.
About the Authors

Lindsay Gorman is the Emerging Technologies Fellow at the German Marshall Fund’s Alliance for Securing Democracy and a consultant for Schmidt Futures. Lindsay has spent over a decade at the intersection of technology development and national security policy, including in the Office of U.S. Senator Mark Warner, the White House Office of Science and Technology Policy, and the National Academy of Sciences. In the latter post, she supported the Committee on International Security and Arms Control in track II nuclear and cyber security dialogues with Chinese and Russian experts. A physicist and computer scientist by training, she previously ran a technology consulting firm, Politech Advisory, advising start-ups and venture capital and has developed cybersecurity tools in Silicon Valley. Her research focuses on understanding and crafting a transatlantic response China’s techno-authoritarian rise, from 5G and the future internet to information manipulation and censorship. Her technical expertise lies in artificial intelligence, statistical machine learning, and quantum materials. Lindsay holds an A.B. in physics from Princeton University, where she graduated magna cum laude, and a M.S. in applied physics from Stanford University.

Bret Schafer is the Alliance for Securing Democracy’s Media and Digital Disinformation Fellow. As an expert in computational propaganda, he has appeared in the New York Times, USA Today, the Wall Street Journal, and the Washington Post, and he has been interviewed on NPR, MSNBC, CNN, Al Jazeera, and CBS and BBC radio. Prior to joining GMF, he spent more than ten years in the television and film industry, including stints at Cartoon Network and as a freelance writer for Warner Brothers. He also worked in Budapest as a radio host, in Berlin as a semi-professional baseball player in Germany’s Bundesliga, and in Moscow as an intern in the Public Affairs Section at the U.S. Embassy in Russia. He has a BS in communications with a major in radio/television/film from Northwestern University, and a master’s in public diplomacy from the University of Southern California, where he was the editor-in-chief of Public Diplomacy Magazine.

Clara Tsao is a non-resident fellow at the Alliance for Securing Democracy. She is an online disinformation expert and a civic tech entrepreneur, who recently co-founded the Trust & Safety Professional Association and the Trust & Safety Foundation to support the global community of professionals who develop and enforce principles and policies that define acceptable behavior and content online. Clara is also a non-resident senior fellow at the Atlantic Council’s Digital Forensic Research Lab. Her previous roles include CTO at the U.S. Department of Homeland Security’s Countering Foreign Influence Task Force and the interagency U.S. Countering Violent Extremism Task Force and Senior Advisor for Emerging Technology at the Cybersecurity Infrastructure Security Agency. She has spent a decade working in the technology industry across global teams at Microsoft, Apple, Sony PlayStation, AT&T, and also as a Google and Mozilla Technology Policy Fellow. Clara is also the Board Chair and President of the White House Presidential Innovation Fellows Foundation and a Senior Advisor at Tech Against Terrorism.

Dipayan Ghosh, Ph.D. is the co-director of the Digital Platforms & Democracy Project at the Shorenstein Center on Media, Politics and Public Policy at the Harvard Kennedy School and faculty at Harvard Law School. A computer scientist by training, Ghosh previously worked at Facebook, where he led strategic efforts to address privacy and security issues. Prior, he was a technology and economic policy advisor at the White House during the Obama administration. Named to the Forbes 30 Under 30, he received a Ph.D. in electrical engineering & computer science from Cornell University, an MBA from the Massachusetts Institute of Technology, and completed post-doctoral work at the University of California, Berkeley.
Endnotes

3 Kerry, Morris, Jr., Chin, Turner Lee, “BRIDGING THE GAPS: A path forward to federal privacy legislation.”
6 5 U.S.C. § 552a, Privacy Act of 1974
7 42 U.S.C. § 201, Health Insurance Portability and Accountability Act
9 Senate Bill 1386 - Personal Information: Privacy, California State Senate, September 26, 2002.
11 “What is GDPR, the EU's new data protection law?,” GDPR.EU, accessed May 13, 2021.
13 Maria Korolov, “California Consumer Privacy Act (CCPA): What you need to know to be compliant,” CSO, July 7, 2020.
24 Jeffrey Knockel, Christopher Parsons, Lotus Ruan, Ruohan Xiong, Jedidiah Crandall, and Ron Deibert, We Chat, They Watch, Citizen Lab, May 7, 2020.
26 Knockel, Parsons, Ruan, Xiong, Crandall, Deibert, We Chat, They Watch.
34 Kerry, Morris, Jr., Chin, Turner Lee, “BRIDGING THE GAPS: A path forward to federal privacy legislation.”
35 Abbott, “Time to Build a National Data Broker Registry.”
37 Doss, Cyber Privacy: Who Has Your Data and Why You Should Care.
49 Public Law No: 116 - 207. Internet of Things Cybersecurity Improvement Act of 2020
54 Office of the Attorney General, “California Consumer Privacy Act (CCPA).”
55 Scott Steinberg, “Cyberattacks now cost companies $200,000 on average, putting many out of business,” CNBC, October 13, 2019.