EXECUTIVE SUMMARY

Data nationalism and data location requirements are on the rise. In late 2019, a Russia-backed initiative to enact a new cybercrime treaty passed in the UN—one that would, in the eyes of Russia, better protect “sovereign” interests and control. This follows a rash of data localization requirements enacted in Russia, India, China, and elsewhere. Even in the United States, Senator Josh Hawley introduced legislation that would require Russian and Chinese-owned companies to store data about Americans in America. The EU, meanwhile, has imposed its own set of transfer restrictions on data, designed to protect the privacy of EU citizens and residents. These are complex moves motivated by a range of different nationalistic, security-based, and economic goals—with potentially profound, and at times unintended, economic, security, and geopolitical costs as well.

This report seeks to detail, break down, and analyze the trends in favor of data nationalism—defined loosely as the effort by nation-states to ensure control over data for a range of normative and security-based reasons. It details the various reasons that states engage in various efforts to exert national control over data—some for surveillance and access reasons, some to push particular data protection and privacy regimes, and some to control the kind of information that flows through one’s borders and/or about one’s citizens and residents. It provides a typology and analysis of the various forms of control, details the costs that arise, and ends with a discussion of some of the efforts at pushback.
INTRODUCTION

Data nationalism and data localization efforts are on the rise. The not-so-long ago vision of a free, open, and globally interconnected internet has been replaced by increasing efforts by nations across the world to assert sovereign controls, and limits, over the data that crosses through and is generated in its borders. An approach that was once denigrated as the province of authoritarianism – contrasted with the U.S. push for “one internet, one global community, and a common body of knowledge that benefits and unites us all” —is now being pushed, to varying degrees, by countries around the world. In the United States, U.S. Senator Josh Hawley (R-Mo.) has introduced legislation that would require Russian- and Chinese-owned companies to store data about Americans in America. Last year, Congress enacted changes to the Committee on Foreign Investment in the United States (CFIUS) authorities so that non-controlling foreign investment in companies that collect or maintain “sensitive” personal data are now subject to national security reviews—and potential blocks on such investment. And there has been a renewed push for export controls on national security-sensitive technology applications, such as in the realm of artificial intelligence, as a means of protecting against nefarious foreign uses.

Meanwhile, a rash of data localization requirements have been enacted in Russia, India, China, and elsewhere that require geographically localized storage and processing of certain kinds of data. Even the EU has imposed its own set of transfer restrictions on data, designed to protect the privacy of EU residents and citizens, but in ways that operate as de facto localization requirements.

Numerous countries also continue to exercise—or at least attempt to exercise—tight controls on the information that crosses their borders. Beijing has long imposed controls on the internet content accessible to residents within China. Russia’s new so-called domestic internet law grants the government significantly greater authority over internet infrastructure in order to better control data flows and possibly cut Russia’s web off from the rest of the world. A range of other countries have taken steps to shut down the internet completely in order to deal with unwanted content; India alone shut down parts of its internet more than any other country in 2018 and again in 2019.

Put simply, once-forceful calls for a “free and open” internet are increasingly being replaced by efforts to protect, control, and manage the internet and its related risks, often in a geographically bordered way. These and many other measures are all classifiable as forms of data nationalism—defined loosely as the effort by nation-states to ensure control over data for a range of security, privacy, normative, and economic-based reasons. The measures are being initiated by countries both democratic and authoritarian; by countries with both advanced economies and emerging market economies; by countries located in the north and south, east and west. And the costs are potentially many: to the economy, to security, to privacy, and to the balance of powers geopolitically.

That said, the purpose, nature, and implications of these moves toward data nationalism vary significantly. Some are done in pursuit of national security and law enforcement objectives. Others are motivated by data privacy and consumer protection concerns. Several are a response to perceived foreign interference threats and flows of unwanted information. Many look to bolster domestic innovation and local industry. In the majority of cases, a combination of motivations overlaps in difficult-to-disentangle ways. The nature and effect of the measures vary significantly as well. Even data localization efforts are not all the same. Some require copies of particular data to be stored locally; others mandate that data cannot leave the jurisdiction at all. Both have costs, but the nature of the costs differs.
At the same time, there also is pushback and internal contradiction among even the most ardent supporters of cyber sovereignty. The U.S. and other countries have criticized data localization requirements as a potential violation of countries’ World Trade Organization (WTO) obligations. Internet freedom advocates, human rights activists, industry groups, and others have criticized efforts to geographically segment the market as diminishing the relative freedom and openness of today’s global internet.

This report details, breaks down, and analyzes the trends in favor of data nationalism. It details the various reasons that states in engage in various efforts to exert national control over data, provides a typology and analysis of the various forms of control, highlights some of the costs, and details as well some efforts at pushback. None of this is simple or one-sided. Even governments like China, long associated with promoting a strong version of data nationalism (also referred to as “cyber sovereignty”) struggle to balance the economic benefits of relatively open internet data flows with the political security and benefits of data control. Conversely, there are those who generally support open data flows online—like India, for example—who have increasingly taken steps to promote local technology firms and data localization requirements in reaction to, among other things, what is perceived as excessive foreign dominance and influence in their domestic technology markets. Understanding the types of, motivations for, and conflicting complex impulses behind data nationalism are the essential first steps in figuring out how best to respond.

**RATIONAL FOR DATA NATIONALISM**

Since the inception of the internet, states have sought to exert control over the data and tech companies that managed the data that flowed through the internet. In the early internet days, this played out primarily via debates over taxation and regulation of e-commerce, but over time these efforts have expanded to include wider-ranging efforts to control the flow of data in and out of a country’s jurisdiction in ways that have led some to worry about the “fracturing” of the global internet—the breaking down of what was once envisioned as a global, open, interconnected communication system into separate, discrete systems.

The motivations behind data nationalism pushes are intersecting and numerous—and their effects varied as well. Here, we break down some of the primary reasons why states are increasingly seeking to exert territorial-based controls over data:

**Table 1**: Rationales For Data Nationalism

<table>
<thead>
<tr>
<th>Type of Rationale</th>
<th>Explanation of Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access to data</strong></td>
<td>Ensuring law enforcement and/or intelligence and security services have the requisite legal and/or technical ability to access data for investigative, prosecutorial, and intelligence reasons</td>
</tr>
<tr>
<td><strong>Consumer protection and privacy</strong></td>
<td>Ensuring the government can protect citizen and resident data against privacy incursions, including surveillance by foreign companies and governments</td>
</tr>
<tr>
<td><strong>Protecting against foreign access</strong></td>
<td>Ensuring that other countries have reduced ability to access citizens’ data for adversarial purposes</td>
</tr>
<tr>
<td><strong>Content controls</strong></td>
<td>Ensuring state authorities have the ability to curate content in accordance with local norms</td>
</tr>
<tr>
<td><strong>Economic</strong></td>
<td>Promoting and protecting the local tech industry</td>
</tr>
</tbody>
</table>
• **Access to Data:** Increasingly, law enforcement finds itself in situations in which data critical to the investigation and prosecution of wholly local crime (e.g., local perpetrators, victims, witnesses) is located outside of its borders. Often, such data is in the hands of tech companies that are prohibited from transferring such data to the investigating state. In order to access that data, law enforcement from the investigating country is required to make mutual legal assistance requests to the nation where the relevant tech provider is located. But that is a time-consuming process. At times, law enforcement may not even know where to direct such a request. Data localization requirements are a means of countering this phenomenon. By requiring the data be held locally, government officials can better ensure access for law enforcement, intelligence, or other reasons.

• **Consumer Protection and Privacy:** Conversely, and especially in the wake of the Edward Snowden revelations regarding the scope of U.S. surveillance activities, governments have sought to secure their citizens’ and residents’ data from what is perceived as excessive surveillance and insufficient privacy protections by foreign states. This is reflected in the European Union’s General Data Protection Regulation (GDPR), which requires that companies handling Europeans’ data do so pursuant to a very detailed set of data protection rules. Companies also are prohibited from transferring Europeans’ data out of Europe absent what is known as an “adequacy” determination—a finding that the destination country meets requisite data protection standards—or alternative mechanisms or agreements that are deemed to provide sufficient protections. Currently, a combination of what are known as standard contractual clauses and a U.S.-EU “Privacy Shield” agreement permits transfers of data to the United States by participating companies, although both mechanisms are currently being challenged in the European Court of Justice as insufficiently protective. A separate U.S.-EU agreement authorizes the law enforcement-to-law enforcement transfer of data.

• **Protecting Against Foreign Access:** Separate and apart from both the law enforcement access and privacy-based interests, states also act in order to address the national security implications of foreign access to what is deemed sensitive data. Thus, many nations require or are considering requiring that certain kinds of government data be held locally. Even the United States, which has generally and often vocally opposed data localization measures adopted by other countries, has Defense Department requirements that cloud contractors working for DoD store Department data within U.S. territory. National security concerns similarly motivated the U.S. Congress to demand additional reviews of foreign interests in companies that manage or hold “sensitive personal data” for national security reasons. And they are also behind the current U.S. move to partially ban the Chinese telecommunications company Huawei from supplying American 5G network technology, and the push to persuade its allies to do the same.

• **Content Controls:** States also seek to impose national-level controls on data in attempt to curate content in accordance with local norms. The most well-known example is the so-called Great Firewall in China, which filters communications content coming in and out of China, using algorithms, machine learning technologies, and manual sorting to censor speech. But China is not alone. Nations across the globe increasingly seek to limit communications content as a means of controlling real and alleged harms. This can take several forms, such as real-time government filtering of internet traffic into the country, or legal requirements that private internet platforms remove content at the government’s behest. Singapore’s 2019 fake news law requires the issuance of correction notices—or takedowns—in response to a government minister’s assertion that information is false and prejudicial to Singapore’s national interests. Germany has a particularly stringent hate speech law that requires companies
to take down any “manifestly unlawful content” within 24 hours. Australia has a law prohibiting “abhorrent violent material” and requires social media companies to keep such content off their sites—or face hefty fines. European law requires search engines to comply with its version of the right to be forgotten—something that is done via what is known geo-blocking content, so that what is accessible via a search in Europe is different than what is accessible elsewhere. An Austrian court recently ruled that a Facebook post criticizing a then-leader of the Green Party for her immigration policy was defamatory and that it and any equivalent commentary had to be kept off the site. These, too, are forms of data nationalism—efforts to control the content of communications in accordance with national norms.

**Economic** Some localization requirements and limits on foreign access are motivated by economic reasons—including, most notably, a desire to provide local jobs and help support the growth of the local tech sector. This is a particularly forceful rationale for the many countries concerned about, and wanting to push back against, what is seen as “data colonialism” of the U.S. tech market. In India, for example, concerns about this Western technological dominance have led some to believe that data localization is a way to work to offset the power imbalance between Indian-incorporated tech firms and multinational giants like Facebook and Google.

**Forms of Data Nationalism**

Just as the motivations behind the efforts to exert local control over data are multiple, the nature of such controls takes many forms. This report focuses on three—data localization mandates; content controls; and foreign access limitations—and describes a range of different approaches within each of these categories as well.

**Table 2** Forms of Data Nationalism

<table>
<thead>
<tr>
<th>Type of Rationale</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data localization mandates</td>
<td>Requiring that certain types of data be stored in a specific geographic area in a specific way</td>
<td>India mandates that payment data is locally stored; Russia mandates social media user data on Russians is locally stored; the US prohibits certain DOD cloud contractors from sending data outside US territory</td>
</tr>
<tr>
<td>Content controls</td>
<td>Controlling and limiting content dissemination online</td>
<td>Vietnam’s cybersecurity law criminalizes a range of critiques of the national government; Australia’s terrorist content law mandates rapid takedown of abhorrent violent material posted on social media; Iran, India, and many other countries shut down the internet in 2019 amidst unrest</td>
</tr>
<tr>
<td>Foreign access limitations</td>
<td>Reducing actual or perceived risks of foreign countries accessing or influencing the collection and storage of citizen data, including through technical or legal means</td>
<td>US, Israel, Russia, and other countries are increasing scrutiny of foreign investments in their tech sectors; the US is considering further export controls to limit sensitive data flowing to China</td>
</tr>
</tbody>
</table>
Data Localization Mandates

Data localization mandates broadly refer to the required storage of certain kinds of data in certain geographic locations. The mandates can differ in scope, substance, and impact depending on what they are trying to achieve.

What we are calling soft data localization involves “mirroring,” which requires that copies of a certain kind of data must be stored within a certain area but does not restrict the copying or transmission of that data elsewhere. For example, a country might require that information on citizens’ social media use must be stored within its geographic borders for various law enforcement or intelligence reasons, while still permitting the social media company to send that same data, or even store that same data, elsewhere in the world. For example, India has mirroring requirements specifically around payment data. Russia has mirroring requirements that require internet sites with data on Russian citizens to store the information within the country. Russia also requires that specific types of data such as encryption keys also must be held locally. But these are just a few examples of many.

What we refer to as hard data localization mandates require that certain data be stored within certain borders and not copied or transmitted anywhere else. This is less common, but increasingly applied in an effort to protect data that is deemed particularly sensitive. The most recent draft of India’s Personal Data Protection Bill contains hard data localization requirements with respect to “critical personal data,” which must be stored and processed only within India. By comparison, “sensitive personal information” is only subject to mirroring requirements.

At times, hard data localization requirements will be conditional or accompanied by waiver provisions. For example, citizen payment data might be subject to hard localization requirements, but with provisions that allow for copies to be transferred abroad under certain specific conditions (i.e., encrypted in a certain way). The EU’s transfer restrictions are an example of conditional hard localization requirements—mandating that EU citizen data be held locally, absent the existence of sufficient data protection provisions to govern the transfer. So too is Article 37 of China’s Cybersecurity Law, which requires that “important” and “personal” data, produced by so-called critical infrastructure operators, must either meet a governmental audit – albeit based on unclear criteria – or be locally stored.

Some governments also have required the establishment of an in-country office as part of data localization requirements. Vietnam, for example, required the opening of an office in-country in its 2019 cybersecurity law, for certain firms collecting data on Vietnamese citizens. These kinds of requirements can be coupled with soft or hard localization requirements, or some requirements in between them on a spectrum of severity.

In addition, and as the foregoing suggests, different kinds of data are often treated differently by the same country, depending on the nature of the data and its perceived sensitivity and importance. Thus, some data may be free of any kinds of localization-based requirements; some other data may be subject to mirroring requirements; and particularly sensitive data could be subject to strict localization rules.
Content Controls

China’s Great Firewall is infamous as one of the most extreme efforts to control what information a country’s citizens and residents can and can’t see on the internet. It operates via filtering of information in and out of the country, although residents traditionally have been able to use a range of technical tools to evade some of these restrictions. More recently, high-profile fights between the Chinese government and the NBA have highlighted again the degree to which Beijing seeks to mute criticism of its policies and practices.  

China, however, is hardly the only nation that seeks to control content accessible to its citizens and residents. A range of countries place limits on what their residents can access online. Russia recently passed a so-called domestic internet law, which grants Russia’s media regulator significantly greater authority over internet infrastructure in order to better control data flows and possibly cut Russia’s web off from the rest of the world. Encrypted messaging app Telegram is banned by law in Russia (though technical enforcements have been largely unsuccessful), as are Virtual Private Networks (VPNs) that could enable circumvention of internet censorship and data flow controls.

In 2019, Vietnam passed a cybersecurity law that criminalized online “propaganda” against the government, as part of a broader mimicry of Beijing’s online control. That same year, Singapore passed a fake news law that permits any government minister to demand a correction notice alongside, or in some cases, takedown of allegedly false and prejudicial information online.

Numerous countries are also using the blunt instrument of internet shutdowns as a means of responding to or preventing the spread of unwanted content and communications. India has executed the most internet shutdowns of any country in the last two years. In January 2020, the Indian Supreme Court ruled that the government’s five-month shutdown of the internet in Kashmir was unlawful, but did not require its immediate restoration. Since then, 2G service has been restored for verified users, but social media remains prohibited and only approved websites can be accessed in the region, according to a group that tracks internet shutdowns in India. In November 2019, Iran imposed a near-total shutdown of the internet as a means of controlling protests over rising gas prices. Many other countries also imposed some kind of temporary network shutdown in 2019, although with significant variances in scope and duration.

Governments also demand that social media companies and other curators of content online abide by their local speech rules. Thailand, for example, prohibits any and all critiques of the Thai monarchy. Turkey bars critiques of Mustafa Kemal Atatürk, the first President of the Republic of Turkey. Germany bans hate speech online, requiring swift action by social media companies to take down content that is deemed to run afoul of this prohibition. Companies that wish to operate in these countries must abide by these rules—or risk being penalized or kicked out. This in turn is its own form of data nationalism—carried out into effect either by local companies that act locally and thus comply with these rules or major international tech companies geo-blocking content in order to comply with local and national laws.

Here, too, these content controls manifest in softer and more stringent versions. In the examples just described, governments generally demand that companies offering services within their territories take action to ensure that residents cannot access the prohibited content—something that large multinational companies often achieve via the use of technical tools. With geo-blocking, companies can offer different user experiences to users located in different places.
So long as users are not relying on VPNs or other tools designed to hide a user’s location, users in a particular location will not be able to access banned information, whereas users located outside that jurisdiction can.

But, with some growing frequency, nations are seeking to impose their content rules globally. This has long been the case with respect to U.S. copyright law; U.S. companies subject to U.S. jurisdiction are obliged to abide by notice-and-takedown requirements, and keep copyright-infringing material off their sites, regardless of where the material is accessed. Additional governments and courts are making similar demands, increasingly in situations in which speech norms in one jurisdiction conflict with those in others. In a high-profile case that worked its way to the European Court of Justice, an Austrian court demanded that Facebook take down and keep off allegedly defamatory content calling the then-leader of the Austrian Green Party a “corrupt oaf,” “lousy traitor,” and “fascist” for her policy stance on immigration. The European Court of Justice ruled that while Austria should be cautious in issuing such kinds of take-down and keep-off orders with global reach, nothing in European law prevented it from doing so. This was one of a handful of cases in recent years raising similar issues.

These are all forms of data nationalism—governments asserting national control and national norms over information that crosses into their borders.

**Foreign Access Limitations**

In addition to the kinds of hard data localization laws that explicitly prohibit transfers of certain data out of the country, governments around the world are taking a range of different steps to limit foreign access to certain kinds of data over which the government has access and control.

In the United States, for example, Congress recently expanded the jurisdiction the Committee on Foreign Investment in the United States—which is authorized to review, and if appropriate, block foreign acquisitions of or investments in American-incorporated companies for reasons pertaining to U.S. national security—to cover non-controlling foreign investment in companies that collect or maintain “sensitive” personal data under its control. This authority was invoked in March 2019 when CFIUS told Chinese company Beijing Kunlan Tech to sell gay dating app Grindr, due to worries about potential Chinese government access to sensitive information on U.S. citizens. In November 2019, CFIUS also reportedly launched a review of the acquisition by Chinese company Bytedance Technology—the parent company of the popular social media app TikTok—of its $1 billion acquisition of U.S. social media app Musical.ly.

The United States has similarly employed and proposed export controls on data-related technologies as a means of limiting foreign access to data. American proposals to limit the spread of machine learning technologies and capabilities abroad, mainly to China, have targeted data-related exports such as training data and deep learning software libraries. This is an example of a targeted export control that is aimed at limiting the spread of certain sensitive data to foreign governments.

In addition, the United States has banned foreign-produced and managed software and hardware systems due to similar concerns about foreign access to its citizens’ information, as well as potentially sensitive corporate or government data. The U.S. Department of Defense banned the use of software made by Kaspersky Labs, the Russian-based antivirus company, because policymakers were concerned it could give the Kremlin backdoor access to government networks. Concerns about Huawei’s
5G technology have similarly led policymakers to limit its presence within American telecommunications infrastructure. Legislation introduced by U.S. Senator Josh Hawley (R-Mo.) in November 2019 combines data localization and foreign access restrictions. It would require Russian- and Chinese-owned companies to store data about Americans in America.\textsuperscript{59}

Other countries are engaging in similar practices that scrutinize and limit both the spread of foreign tech companies and potential foreign government access to data on their citizens.

Beijing has long put barriers in the way of foreign, especially American, tech companies—often by imposing excessive censorship requirements and refusing to let in or kicking out companies that fail to comply, even if they are not imposing outright bans. Facebook, for example, has engaged in multiple—ultimately unsuccessful—efforts to gain access to the Chinese markets. Mandatory source code inspections also have presented difficulties for American tech firms.\textsuperscript{60} The recent introduction of hundreds of cybersecurity standards has likewise made the market an increasingly difficult place for foreign companies to operate due to heavy compliance costs.\textsuperscript{61}

Moscow too has increasingly taken steps to limit what is perceived foreign influence. Its fears about foreign influence over its domestic technological sphere grew notably during the Arab Spring and have since manifested in various policy pushes to limit foreign access to and influence over Russian data and technology systems.\textsuperscript{62} The Kremlin, for instance, recently backed a bill limiting foreign ownership of Russian-incorporated technology firms. This led to investors bailing on shares of Yandex, the Russian internet service provider—something that Yandex warned would hurt its efforts to expand internationally.\textsuperscript{63} The Russian government also last year moved to replace the Microsoft Windows operating system on government computers due to concerns about Western espionage.\textsuperscript{64} Most recently, Russian President Vladimir Putin recommended that legislation to protect foreign investments in Russian companies against security risks be developed and approved by April 30, 2020.\textsuperscript{65} Reducing technological reliance on the West, and reducing the West’s influence on Russia’s domestic technology sphere, underpin these actions.

The Indian military similarly has prohibited military personnel from installing Chinese social messaging app WeChat on their phones over similar national security concerns.\textsuperscript{66} These are all forms of data nationalism—motivated by a combination of concerns about foreign access to potentially sensitive data and a desire to promote, via the exclusion of others, local tech products and companies.

\section*{Potential Costs of Data Nationalism Efforts}

Data nationalism is, in some ways, an inevitable and often at times fully appropriate means of national governments exerting control over the data that flows through and companies that operate in their jurisdictions. Is it part and parcel of an effort to protect their own citizens and residents from what is perceived as malicious interference by others, pursue economic growth and strength, ensure access to data to investigate and prosecute local crime, and maintain and assert a normative vision of what is and is not permitted speech.

But at the same time, many of the moves toward stronger data nationalism have a range of potential side-effects and costs. The following describes an array of different economic, knowledge-based, rights-based, and security costs that can and do result—although the scale and scope of such costs obviously varies based on the nature and extent of the measure(s) being imposed.
Table 3: Potential costs of data nationalism

<table>
<thead>
<tr>
<th>Type of Cost</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>Costs to companies operating multinationally—buying servers; changing data storage and routing protocols; additional cybersecurity protections</td>
</tr>
<tr>
<td>Knowledge-based</td>
<td>Cutting off valuable tech-related information transfers across borders</td>
</tr>
<tr>
<td>Rights-based</td>
<td>Facilitating access to citizen data for censorship or surveillance purposes</td>
</tr>
<tr>
<td>Security</td>
<td>Risk that local storage with weaker security measures potentially undermining data security</td>
</tr>
</tbody>
</table>

- **Economic**: Data localization requirements in particular can impose significant financial costs on companies that seek to do businesses across multiple different countries. The costs of building or leasing server space and ensuring secure storage are, according to some analyses, high. Additional costs may result from altering back-end data-routing infrastructure (such as modifying cloud storage protocols) in order to comply with data localization laws that vary in terms of severity and specific technical requirements. Localization requirements also can impact network efficiency. Generally, systems that keep data closer to a user can deliver that data more quickly to the user seeking access. But localization requirements also could impose latency costs—in other words, cause slight network delays—for how the company routes data to other servers around the world, which may in turn impact competitiveness. Country-by-country content-based rules also impose their own set of economic costs by requiring companies to set up multiple different content-review policies and procedures, which could require additional technical platform alterations. Doing this country-by-country content review well also requires the hiring of content-curators who speak a local language and understand the local culture, so as to be able to distinguish between, say, hate speech and parody. Many will note that these are costs that major, multinational, multi-billion dollar companies can and must take on as a condition of accessing new markets. But there also is a risk that these kinds of costly requirements end up entrenching the power of the small few that can actually afford to meet the multiple requirements—and limiting the ability of smaller players, including the local tech industries that many national governments are eager to support, to compete on a global scale.

Export controls similarly also impose costs of compliance—costs that may be well worth it (at least from a state’s view) depending on the nature of the control. These costs can increase as controls are insufficiently clear and well-defined—making it both more expensive and riskier for companies to grow into foreign markets. Overall, specific costs vary depending on the specifics of an export control policy, and costs may be more predictable where controls are defined with an appropriate level of specificity.

- **Knowledge-Based**: Research often requires access to large data sets. Development of precise and accurate AI systems that works well across a variety of demographics often depends (though not exclusively) on the ability to gather and use large data sets with an array of different inputs, including from multiple different jurisdictions. Those developing human language translation programs, for example, require access to what is often foreign-based data. Transfer restrictions can thus limit or curtail valuable research. U.S.-proposed export controls on national security-
sensitive technologies, for example, were widely criticized in this vein as overly broad, poorly defined, and likely to unnecessarily curb the benefits of global and open AI research.\textsuperscript{69}

\textbf{Rights-Based}: Local control can be a means of ensuring compliance with domestic norms. But it can also be used by authoritarian regimes to stifle dissent, engage in excessive surveillance, and engage in abusive means of digital control. While, for example, there is a need for multinational companies to respect local laws and divergent norms across borders, there also is a risk that data nationalism can be used as an excuse for local censorship and repression. When Russia successfully got UN approval to move forward with new cybercrime treaty in 2019—one that is in the eyes of the Russians meant to better reflect sovereign interests than the Council of Europe's competing Budapest Convention—several commentators rightfully noted the risk that it could be used by authoritarian regimes to provide political cover for prosecution of political opponents.\textsuperscript{70} This is just one example among many.

\textbf{Security}: Localization requirements also risk undermining cybersecurity. Mirroring, through requiring additional storage of a copy of certain kinds of data in a certain geography, can mean that there is more redundancy in the system and thus more means of recovering data in the case of some sort of attack. But mirroring also increases the attack surface to access that data—a particular concern if mandated in a location that fails to apply high security standards to data at-rest. Localization can also impede analysis of data patterns that is used to detect things like money laundering and other kinds of malicious online activity.\textsuperscript{71} In addition, the notion that the local storage of data makes it better protected against access by foreign governments is not necessarily true. If data is stored locally with weaker security protections, it could undermine that data's protection.\textsuperscript{72} Moreover, and perhaps ironically, several nations are much more permissive in terms of the kind of surveillance—and the ex-ante review and ex post oversight that occurs—with respect to collection of data located outside their territory as compared to surveillance of data that is territorially located.\textsuperscript{73}

\section*{PUSHBACK}

Notably, the trends in favor of data nationalism do not all push in the same directions. Powerful, multinational companies are often effective in pushing back against data policies that impose costs on them—such as financial costs (i.e., buying servers), technological costs of locally storing certain kinds of information (i.e., potential traffic latency), or reputational costs of having to comply with government censorship or data access requests. Success of these pushback efforts depend on a number of factors, such as the political and financial influence of a given firm and the economic leverage wielded by a government (i.e., how valuable it is for the company to access the market).

The U.S. government has also supported such pushback—at least when the restrictions come from abroad. U.S. pushback has, for example, been widely cited as leading India to relax some of the hard data localization requirements in the most recent draft of its Data Protection legislation.\textsuperscript{74} Japan, particularly Japanese Prime Minister Shinzo Abe, also has been outspoken in promoting efforts to advance global data governance. The phrase “data free flow with trust” has been the government’s way of describing how countries should promote relatively free data flows with appropriate protections in place.\textsuperscript{75} That said, India’s refusal to sign the digital economy declaration out of the 2019 G20 in Osaka—whose language was principally shaped by the U.S. and Japan—underscores the potentially limited influence of these efforts.\textsuperscript{76}
Pushback may also originate from civil society. Multiple different groups monitor and seek to put pressure on governments that engage in what is perceived as excessive censorship and surveillance—putting pressure as well on companies that might otherwise comply. Content takedown laws in Australia, India, Germany, and many other countries likewise have received criticism from civil society. Foreign access limitations proposed in the U.S. in the form of export controls were criticized by a variety of industry groups and technology policy experts for harming industry and diminishing the benefits of collaborative AI research.

In China’s case, however, companies tend to comply with government policies or not operate in the market at all. Facebook was kicked out of China (not for the first time) in 2018 apparently due to government problems with its lack of censorship on the platform; Google confirmed its termination of a censored search engine in China in 2019 after pushback from employees and the American media, and criticism from U.S. officials; the list goes on. All to say, no company has successfully pushed the Chinese government to reduce data nationalism policies, from content controls to data localization requirements. Companies either comply or leave.

**CONCLUSION**

Data nationalism is on the rise. The once heralded vision of a free, open, and globally interconnected internet has been replaced by an increasingly strong push to demand data be stored locally, erect virtual borders, and keep foreign influence out—motivated by an array of economic, normative, and security-related costs. Many are based on the very legitimate interests of states to ensure access to data for the investigation of criminal activity; some motivated by an interest in protecting citizens’ and residents’ privacy; and others an attempt to support of local industry and keep foreign competitors out. Such moves can protect local norms and institutions; conversely, they can enhance the ability to spy on citizens and suppress free speech. The ultimate goal is to identify and address the legitimate state interests while also protecting and promoting the many benefits that arise from the cross-border flows of data and information.
ENDNOTES

1 Jennifer Daskal is a Professor and Faculty Director of the Technology, Law & Security Program at American University Washington College of Law. Justin Sherman is a Nonresident Fellow at the Atlantic Council’s Cyber Statecraft Initiative and a Fellow at the Duke Center on Law & Technology at Duke University School of Law.


15 See, e.g, Beina Xu and Eleanor Albert, Media Censorship in China (Washington, D.C.: Council on Foreign Relations, February 17, 2017), https://www.cfr.org/backgrounder/media-censorship-china ("The Chinese government has long kept tight reins on both traditional and new media to avoid potential subversion of its authority. Its tactics often entail strict media controls using monitoring systems and firewalls, shuttering publications or websites, and jailing dissident journalists, bloggers, and activists... At the same time, the country’s burgeoning economy relies on the web for growth, and experts say the growing need for internet freedom is testing the regime’s control").


18Cybercrime Convention Committee, Criminal justice access to electronic evidence in the cloud: Recommendations for consideration by the T-CY (Strasbourg: Council of Europe, September 16, 2016), 15. https://rm.coe.int/16806a495e.

19General Data Protection Regulation, arts. 44-50.

20Data Protection Commissioner vs. Facebook Ireland Limited, Maximillian Schrems (2018), Reference for a preliminary ruling from the High Court (Ireland) made on 9 May 2018 (High Court Ireland), http://curia.europa.eu/juris/document/


23This regime continues to evolve today, such as through establishing a state-approved Virtual Private Network that the Chinese government can use to collect even more information on citizens and possibly execute additional modifications of information accessible to those within China. See Peter Hansen’s commentary in Jordan Schneider’s ChinaEconTalk newsletter: “A Legal Chinese VPN? The Next Phase in Internet Censorship,” ChinaEconTalk Newsletter, January 6, 2020, https://chinaecontalk.substack.com/p/a-legal-chinese-vpn-the-next-phase.


35 “The Personal Data Protection Bill, 2019,” https://drive.google.com/file/d/1vmeCRehq7eiURatOhnio_UTaCkSgM5gv/view  
36 These kinds of requirements also have the perhaps unintended consequences or pushing companies to scan user data to classify data types—and determine whether it is the kind of either “sensitive” or “critical” data subject to localization requirements. This could ironically generate privacy issues directly out of a bill titled around personal data protection.  
45 https://internetshutdowns.in/.  
51 Case C-18/18, Glawischnig-Piesczek v. Facebook Ireland Ltd. [Judgment Facebook Opinion], ECLI:EU:C:2019:821, ¶ 12 (Oct. 3, 2019).


The United States for example imposes numerous procedural and substantive limitations on the foreign intelligence collection of territorially-located data that do not apply if and when there is equivalent collection of extraterritorially-located data.


See a transcript of Prime Minister Abe’s speech at the January 2019 Davos summit, a foreshadowing of more work to come at the G20 in Osaka, Japan later that year: Shinzo Abe, “Defeatism about Japan is now defeated,” World Economic Forum, January 23, 2019, https://www.weforum.org/agenda/2019/01/abes-speech-transcript/.


