Atlantic Council
SCOWCROFT CENTER
FOR STRATEGY AND SECURITY

Atlantic Council
EUROPE CENTER

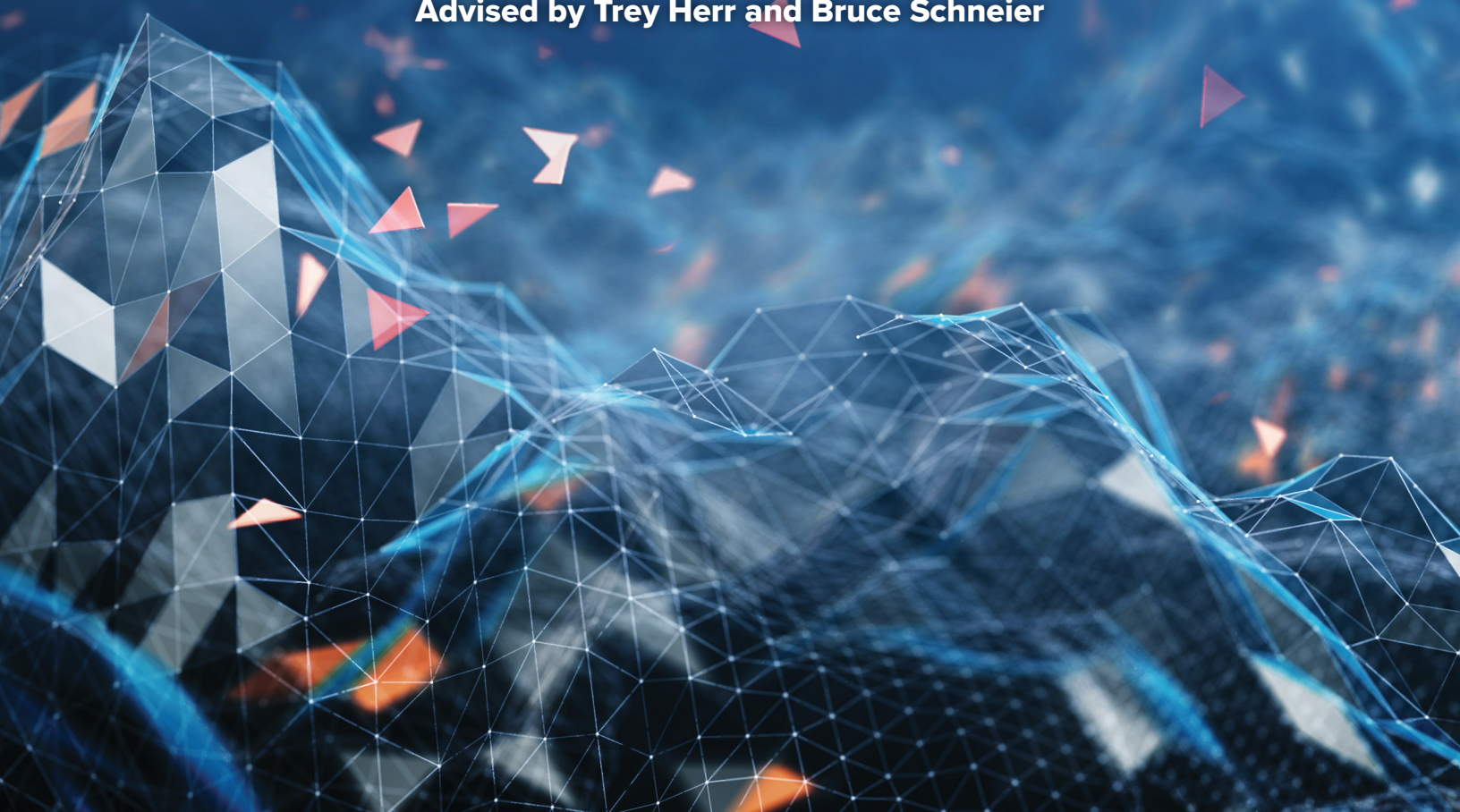
CYBER STATECRAFT
INITIATIVE

Data Rules for Machine Learning:

How Europe can Unlock the Potential While Mitigating the Risks

Blanka Soulava, Hamish Cameron and Victoria Ying

Advised by Trey Herr and Bruce Schneier



The **Scowcroft Center for Strategy and Security** works to develop sustainable, nonpartisan strategies to address the most important security challenges facing the United States and the world. The Center honors General Brent Scowcroft's legacy of service and embodies his ethos of nonpartisan commitment to the cause of security, support for US leadership in cooperation with allies and partners, and dedication to the mentorship of the next generation of leaders.

The **Cyber Statecraft Initiative** works at the nexus of geopolitics and cybersecurity to craft strategies to help shape the conduct of statecraft and to better inform and secure users of technology. This work extends through the competition of state and non-state actors, the security of the internet and computing systems, the safety of operational technology and physical systems, and the communities of cyberspace. The Initiative convenes a diverse network of passionate and knowledgeable contributors, bridging the gap among technical, policy, and user communities.

The **Europe Center** conducts research and uses real-time commentary and analysis to guide the actions and strategy of key transatlantic decisionmakers on the issues that will shape the future of the transatlantic relationship and convenes US and European leaders through public events and workshops to promote dialogue and to bolster the transatlantic partnership.

The Atlantic Council's **Transatlantic Digital Marketplace Initiative** seeks to foster greater US-EU understanding and collaboration on digital policy matters and makes recommendations for building cooperation and ameliorating differences in this fast-growing area of the transatlantic economy.

Atlantic Council
SCOWCROFT CENTER
FOR STRATEGY AND SECURITY

Atlantic Council
EUROPE CENTER

CYBER STATECRAFT
INITIATIVE

Data Rules for Machine Learning:

How Europe can Unlock the Potential While Mitigating the Risks

Blanka Soulava, Hamish Cameron and Victoria Ying

Advised by Trey Herr and Bruce Schneier

ISBN-13: 978-1-61977-197-0

Cover image: iStock/DKosig

November 2021

Table of Contents

Executive Summary

Artificial intelligence (AI) will increasingly shape societies and the global economy. Machine learning—which is responsible for the vast majority of AI advancements—is enhancing the way businesses and governments make decisions, develop products, and deliver services.[1] For machine learning algorithms to learn and make increasingly accurate predictions, they need large quantities of high-quality data that is relevant, representative, free of errors, and complete.[2] As such, data policies have a significant bearing on an economy's capacity to take advantage of machine learning.

Supremacy in AI technologies has become a key aspect of strategic competition between China and the United States. President Xi Jinping has made achieving global leadership in AI by 2030 central to building China into a "modern socialist power."[3] China's rapid advances in machine learning have alarmed America's private sector and national security establishment.[4] The Biden administration is looking to shore-up America's national footing to compete with China in an AI arms race.[5] This competition also extends to setting the global rules and norms that govern AI systems, given their influence on how economies and societies function.[6]

The European Union (EU) already lags behind China and the United States in key AI indicators such as private investment,[7] patent filings,[8] and data market growth.[9] As China and the United States move to position their laws, bureaucratic structures, and public resources to accelerate the development and deployment of AI technologies,[10] the EU faces the added challenge of coordination among its twenty-seven member states to do the same.

Aware of this challenge, the European Commission has made "a Europe fit for the digital age" a priority to achieve by 2024.[11] It is pursuing a "third way" between China and the United States in shaping its digital future, particularly when it comes to the use of data and AI.[12] This has included a raft of legislative proposals, with more in the pipeline, informed by the commission's *European Strategy for Data* and its *White Paper on Artificial Intelligence*, both released in 2020.[13] Those proposals grapple with striking the right balance between reaping the economic benefits of those technologies while minimizing security threats

1 Karen Hao, "What Is Machine Learning?," *MIT Technology Review*, November 17, 2018, accessed July 25, 2021, https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/.

2 European Commission, Proposal for a Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Forwarded to the European Parliament and the Council, 2021/0106/COD, April 21, 2021, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence.

3 Xi Jinping, Remarks at the Nineteenth National Congress of the Communist Party of China (CCP) on October 18, 2017, full text of report via Xinhua News Agency, China Daily (website of CCP English-language news organization), accessed May 27, 2021, https://www.chinadaily.com.cn/china/19thcpcnationalcongress/2017-11/04/content_34115212.htm.

4 Graham Allison and Eric Schmidt, Is *China Beating the U.S. to AI Supremacy?*, Harvard Kennedy School, 2020, 34.

5 Robert O. Work, Remarks of the Vice Chair, National Security Commission on Artificial Intelligence, Press Briefing on Artificial Intelligence, United States Department of Defense (website), April 9, 2021, https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/2567848/honorable-robert-o-work-vice-chair-national-security-commission-on-artificial-i/.

6 Justin Sherman, *Essay: Reframing the U.S.- China AI "Arms Race*," New America, March 6, 2019, https://www.newamerica.org/cybersecurity-initiative/reports/essay-reframing-the-us-china-ai-arms-race/.

7 Frances G. Burwell and Kenneth Propp, *The European Union and the Search for Digital Sovereignty: Building "Fortress Europe" or Preparing for a New World?*, Atlantic Council, June 2020, https://www.atlanticcouncil.org/wp-content/uploads/2020/06/The-European-Union-and-the-Search-for-Digital-Sovereignty-Building-Fortress-Europe-or-Preparing-for-a-New-World.pdf.

8 World Intellectual Property Organization, "Technology Trends 2019 Executive Summary: Artificial Intelligence," accessed May 27, 2021, https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055_exec_summary.pdf.

9 "The European Data Market Study Update," European Commission (website), accessed 2021, https://digital-strategy.ec.europa.eu/en/library/european-data-market-study-update.

10 The White House, "Fact Sheet: The American Jobs Plan," White House (Briefing Room website), March 31, 2021, https://www.whitehouse.gov/briefing-room/statements-releases/2021/03/31/fact-sheet-the-american-jobs-plan/; and Jinping, Remarks at the Nineteenth National Congress of the CCP.

11 European Commission, "6 Commission Priorities for 2019-24," European Commission (website), https://ec.europa.eu/info/strategy/priorities-2019-2024_en.

12 Luis Viegas Cardoso, "Panel Discussion during AI, China, and the Global Quest for Digital Sovereignty – Report Launch," GeoTech Center, Atlantic Council, January 13, 2021, https://www.atlanticcouncil.org/blogs/geotech-cues/event-recap-the-global-quest-for-digital-sovereignty-report-launch/.

13 European Commission, *European Strategy for Data*, February 19, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN ; and European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*, 2020, https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

and protecting the public from harm, including preserving individual rights such as privacy and nondiscrimination.[14]

With the digital economy predicted to be valued at $11 trillion by 2025 on one hand,[15] and the risk machine learning can pose to fundamental rights and European autonomy on the other, the cost of getting that balance wrong is high. The European Commission will need to walk a fine line that separates the legitimate interests of businesses, citizens, and national governments to capture the economic benefits while limiting the negative social impacts of machine learning. Doing so will place the EU in a strong position to shape global rules at the intersection of technology and society to reflect European systems and values more closely.

Given the crucial role data plays in developing innovative AI systems, this report focuses on data governance in the context of machine learning. It provides European decision makers with a set of actions to unlock the potential of AI, while mitigating the risks. Developed at the request of the Atlantic Council, a nonpartisan organization based in Washington, DC, it examines the geopolitical considerations of the EU in the context of competition with China and the United States, the domestic considerations of EU member states including societal interests and the protection of individual rights, and the commercial considerations of EU-based businesses.

The report builds on the existing regulatory initiatives of the European Commission and attempts to lay out key considerations of the EU, the member states, and businesses. In doing so, we believe it captures the dominant features of the current policy debate in Europe to help navigate the complex equities at stake in developing data rules. That said, framing the issue of data governance around international political and economic competition can result in lines of reasoning that pay less consideration to important societal interests and can trivialize complex and multidimensional issues, such as social equity and individual rights. Therefore, we devoted the last section of this report to the protection of fundamental rights, with two chapters focused on these issues. We also have attempted to flag these cases in other parts of the report when they arise in our analysis.

Summary of Recommendations

Leverage the Scale of Europe

Addressing fragmented open data rules and procedures across Europe and harmonizing technical data standards would increase the quantity of data available for innovation and reduce the complexity and cost of preparing data for machine learning.

Therefore, the European Commission should:

1.1A Establish an EU data service under legislation creating common data spaces to facilitate access to public-sector data in consultation with national data-protection authorities by:

i. Serving as a single window to grant access to data sets spanning multiple jurisdictions under a single or compatible format and license.

ii. Making recommendations to European public bodies on conditional access to sensitive data in consultation with national data authorities.

1.2A Establish an EU technical data standards registry to increase transparency and facilitate sharing of various standards used by EU-based organizations and help identify opportunities to increase interoperability.

1.2B Establish sector-specific expert committees to recommend preferred technical data standards, including:

i. Standard data models that enable the translation of differently structured data across systems and organizations.

ii. Standard application programming interfaces (APIs) that enable apps and databases to seamlessly exchange data.

iii. Free open source data management-software that embed sector-specific standards.

14 Article 2 of the Treaty on European Union states that the Union is "founded on the values of respect for human dignity, freedom, democracy, equality, rule of law and respect for human rights"; see Consolidated Version of the Treaty of the European Union, *Official Journal of the European* Union, 2016/C 202/01 (2016); 20, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12016ME/TXT&from=EN). These values were reinforced by the Treaty of Lisbon with an explicit reference to the Charter of Fundamental Rights of the European Union in Article 6(1); see European Parliament, "The Treaty of Lisbon, Fact Sheets on the European Union," European Parliament (website), accessed August 5, 2021, https://www.europarl.europa.eu/factsheets/en/sheet/5/the-treaty-of-lisbon. The charter recognizes the right to respect for private and family life (Article 7), protection of personal data (Article 8), and nondiscrimination (Article 21) among other rights; see Charter of Fundamental Rights of the European Union, *Official Journal of the European Union (*June 7, 2016), https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12016P/TXT&from=EN).

15 James Manyika et al., "Unlocking the Potential of the Internet of Things," Executive Summary, McKinsey Global Institute, June 1, 2015.

Balance Economic Autonomy and Openness

By promoting diversification in cloud data storage and processing services, the EU can mitigate vulnerabilities to disruptions and coercion. However, it also needs to maintain competitive cloud offerings for European businesses and avoid succumbing to calls for protectionist measures. Simplifying rules and minimizing legal ambiguity around data transfers can also lower the risk to European businesses of being excluded from global value chains, as can removing anticompetitive barriers to accessing data while maintaining incentives for businesses to invest in high-quality data sets.

Therefore, the European Commission should:

2.1A Enhance the implementation of the European Alliance for Industrial Data, Edge and Cloud by seeking additional investments from participating cloud computing providers and industrial cloud users.

2.1B Ensure coordination between the European Alliance for Industrial Data, Edge and Cloud and parallel initiatives and ensure that the emerging legislative proposals do not impose significant regulatory burdens on cloud service providers that could impact the competitiveness of their offerings.

2.2A Clarify legal ambiguity on liability and ownership when transferring data, including in the General Product Safety Directive, Product Liability Directive, and in planned legislation on sector-specific common data spaces.

2.2B Prioritize data protection adequacy negotiations with global innovation hubs that share European values on the protection of personal data to enable the transfer of data outside the EU, including exploring solutions with the United States following the Schrems II decision.

2.3A In planned sector-specific "common data spaces" legislation, adopt sector-specific data portability rights for customer data (personal and nonpersonal) that consider the distinct data access problems at an industry level.

2.3B In planned sector-specific common data spaces legislation, adopt sector-specific data access obligations for platforms and dominant companies to share data with other businesses that consider the business models, public interest, and economic sensitivities of each sector.

2.3C Establish a consultative mechanism involving industry, national standards organizations, the proposed European Data Innovation Board, and competition authorities to periodically review the suitability of sectoral data-access rules as technologies and market structures change.

Protect Fundamental Rights

The EU can ensure the protection of fundamental rights while enabling data use for innovation and protecting privacy by decreasing compliance costs and legal uncertainty of data protection while investing in research on machine learning techniques that reduce the need for large pools of personal data. The EU can also mitigate data-driven discrimination through machine learning by addressing gaps in data protection and antidiscrimination laws and strengthening enforcement capabilities.

Therefore, the European Commission should:

3.1A Coordinate with the European Data Protection Board and national data-protection authorities to provide more detailed guidance and assistance on applying the General Data Protection Regulation (GDPR) to machine learning applications to enhance compliance and mitigate business costs.

3.1B Enhance investment in research in privacy-preserving techniques that can adequately protect personal data while enabling machine learning development, including in partnership with like-minded countries such as the United States.

3.2A Work with the European Parliament and member states on providing more specific requirements in the new EU AI legislation for data-management standards to mitigate risks of discrimination. These should focus on greater transparency in access to AI systems documentation and ensuring that data requirements also address cases where the data sets could embed and amplify discrimination.

3.2B Provide sufficient financial and human resources to national data-protection authorities and other relevant bodies that investigate antidiscrimination violations and carefully define and coordinate the establishment of new enforcement authorities with the existing ones.

3.2C Increase transatlantic cooperation to close the gap between US and EU approaches to mitigating risks of bias and discrimination enabled by data in machine learning through the EU-US Trade and Technology Council.

Introduction

Machine learning is changing the way public- and private-sector organizations operate, make decisions, develop products, and deliver services. It is already a big part of people's daily lives, with a rising number of applications emerging across society from healthcare to transportation. As the technology develops, societies at the forefront of machine learning will become more innovative and their industries more competitive.[16] Those that fall behind risk moving down the value chain and taking a smaller slice of the global economy. However, those that that do not manage data rules and machine learning applications with care risk undermining individual rights and exacerbating existing social problems.

As a major subset of artificial intelligence (AI), machine learning is a powerful tool for analyzing very large data sets and spotting patterns in that data.[17] It has two key ingredients: algorithms that learn and data to train machine learning systems. Self-learning machines take each additional data point and adapt the new information to make more accurate predictions. To make very precise predictions, like diagnosing a disease from an eye scan, machine learning requires very large quantities of reliable data, in this case, accurately labeled eye scans.

Rules that govern data have a significant bearing on an economy's capacity to take advantage of machine learning. Industries that can more easily exploit large quantities of high-quality data will be better able to take advantage of the technology. Autonomous vehicles are a good example of an emerging disruptive innovation that depends on gathering high-quality data to feed algorithms. When large amounts of data are available, data scientists can train autonomous vehicles to be less likely to have a collision. However, the quality of these data sets erodes when data is missing, incomplete, inaccurate, duplicated, or dated.[18]

Importantly, data rules also have a significant bearing on society and individual rights. Rules that govern who can access, collect, use, and store data—such as who you are, where you go, what you say or do and with whom, the products you buy or the maladies you suffer—dictates who gets to surveil your life and on what grounds.[19] As such, data rules can have significant implications on rights widely viewed in Europe as fundamental to freedom and equality, such as the right to respect for private life, the protection of personal data, and the right to nondiscrimination.[20]

Given the economic, social and ideological implications, data rules can exert geopolitical influence on how economies and societies function. Whose rules become more commonly adopted by other countries underpins whose economic, social, and ideological systems dominate global digital norms. For example, how open an economy is to international trade and foreign investment,[21] how free society is from repression,[22] and how protected individuals are from harm.[23]

Leading the development of frontier technologies has become central to the intensifying rivalry between China and the United States, and AI is a key aspect of that competition.[24] Chinese leaders believe that being at the forefront of AI technology is critical for China's economic and national security.[25] President Xi Jinping has outlined policies for China to be the world leading AI power by 2030.[26] This

16 James Manyika et al., "Unlocking the Potential of the Internet of Things," McKinsey Global Institute, June 1, 2015.

17 Ben Buchanan and Taylor Miller, *Machine Learning for Policymakers*, White Paper, Harvard Kennedy School's Belfer Center, 2017, https://www.belfercenter.org/sites/default/files/files/publication/MachineLearningforPolicymakers.pdf.

18 Venkat N Gudivada, Amy Apon, and Junhua Ding, "Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations," in *International Journal on Advances in Software* 10, no.1 (2017): 20.

19 Louis Menand, "Why Do We Care So Much About Privacy?," *New Yorker*, June 11, 2018, https://www.newyorker.com/magazine/2018/06/18/why-do-we-care-so-much-about-privacy.

20 European Union, Charter of Fundamental Rights of the European Union, *Official Journal of the European Union*.

21 Magnus Rentzhog and Henrik Jonströmer, *No Transfer, No Trade: The Importance of Cross-Border Data Transfers for Companies Based in Sweden*, Kommerskollegium (Sweden's National Board of Trade), January 2014, https://www.kommerskollegium.se/globalassets/publikationer/rapporter/2016-och-aldre/no_transfer_no_trade_webb.pdf.

22 Adrian Shahbaz, "The Rise of Digital Authoritarianism," Freedom House (website), 2018, https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism.

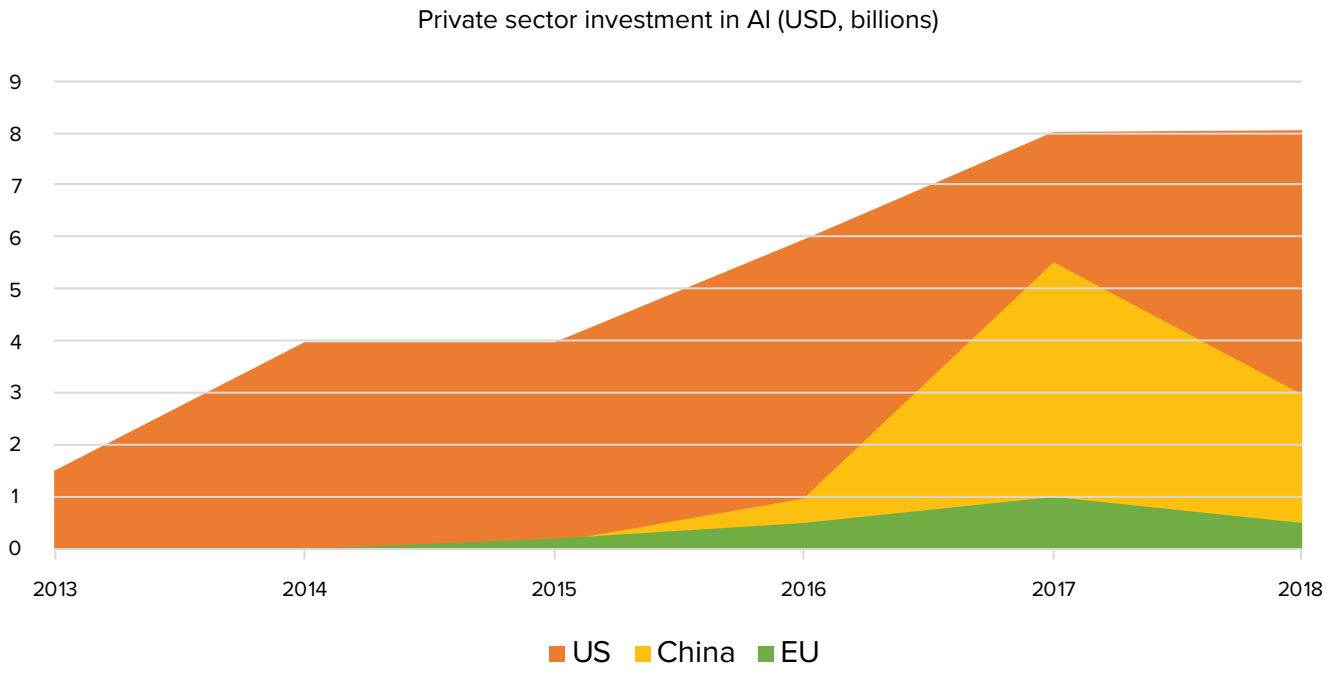
23 European Union Agency for Fundamental Rights, *Getting the Future Right: Artificial Intelligence and Fundamental Rights*, Publications Office of the European Union, 2020, doi:10.2811/774118.

24 Robert O. Work, Remarks of the Vice Chair, National Security Commission on Artificial Intelligence, Press Briefing on Artificial Intelligence, United States Department of Defense (website), April 9, 2021, https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/2567848/honorable-robert-o-work-vice-chair-national-security-commission-on-artificial-i/.

25 Gregory C. Allen, "Understanding China's AI Strategy," Center for New American Security, accessed May 27, 2021, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

26 Ryan Hass and Zach Balin, "US-China Relations in the Age of Artificial Intelligence," Brookings Institution, January 10, 2019, https://www.brookings.edu/research/us-china-relations-in-the-age-of-artificial-intelligence/.

Figure 1: EU Is Outspent on Private Sector Investment in AI[1]



Source: Organization for Economic Co-operation and Development

1 Frances G. Burwell and Kenneth Propp, *The European Union and the Search for Digital Sovereignty: Building "Fortress Europe" or Preparing for a New World?*, Atlantic Council, June 2020, https://www.atlanticcouncil.org/wp-content/uploads/2020/06/The-European-Union-and-the-Search-for-Digital-Sovereignty-Building-Fortress-Europe-or-Preparing-for-a-New-World.pdf.

has included designating national champions that are each focused on different applications of machine learning, like facial recognition and language processing, as well as incentives for start-up investments and patents applications in AI technologies.[27] China now outpaces the United States in AI start-up investment, patents, and performance in international AI competitions.[28] In 2020, China's share of global AI journal citations were higher than the United States for the first time.[29]

In the United States, which has been leading the world in AI,[30] there is bipartisan support in Congress to approve the largest boost to R&D funding in a quarter century,[31] with specific reference to maintaining leadership in AI technologies.[32] The 2021 National Security Commission on Artificial Intelligence Final Report has also called for a new whole-of-government organizational structure to coordinate, incentivize and resource the accelerated wide-scale adoption of AI technologies in the United States—the goal is to position the nation to compete with China.[33]

The European Union (EU) is falling behind China and the United States in machine learning.[34] One telling indicator is the level of private-sector investment in AI, with investment in China and the United States being much larger than the amount seen in the EU, as shown in figure 1. The

27 Allison and Schmidt, *Is China Beating the U.S. to AI Supremacy?*, Harvard Kennedy School, 5.

28 Allison and Schmidt, *Is China Beating the U.S. to AI Supremacy?*, 5–6.

29 Daniel Zhang, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, *The AI Index 2021 Annual Report*, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, March 2021, https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Master.pdf.

30 Daniel Castro and Michael McLaughlin, *Who Is Winning the AI Race: China, The EU, or the United States?— 2021 Update*, Center for Data Innovation, January 2021, 49, https://www2.datainnovation.org/2021-china-eu-us-ai.pdf.

31 John D. McKinnon, "House Passes Bipartisan Bill to Boost Scientific Competitiveness, Following Senate," *Wall Street Journal*, June 29, 2021, https://www.wsj.com/articles/house-passes-bipartisan-bill-to-boost-scientific-competitiveness-following-senate-11624941848.

32 The White House, "Fact Sheet: The American Jobs Plan."

33 Robert O. Work, Remarks of the Vice Chair, National Security Commission on Artificial Intelligence.

34 Castro and McLaughlin, *Who Is Winning the AI Race*.

key challenge the EU faces includes the complexity of coordinating data rules and standards among twenty-seven member states and their fragmented implementation at the national and subnational level. Higher public expectations around privacy and data protection also means the EU must proceed with greater care in making data available for businesses and researchers, slowing the development of machine learning applications in the shorter term, though potentially creating a more sustainable approach that places European innovation on stronger footing in the longer term.[35]

In comparison, China and the United States both have larger, more homogenous data ecosystems that can be more easily accessed by data scientists, better facilitating the training of machine learning algorithms. As a result, the US data market is three times the size,[36] compared to the EU's, and is growing at twice the rate.[37] China's data market is larger still and growing at three times the rate of the EU's.[38] On its current trajectory, the EU will fall even further behind China and the United States, potentially leaving it dependent on foreign supplies of AI technologies trained on data outside Europe.

In an effort to compete globally, the EU is grappling with data policies that strike the right balance between innovation, security, and the protection of fundamental European rights. Privacy and security concerns in Europe have already undermined public trust in foreign governments and private companies. These concerns are amplified by Europe's data infrastructure being largely controlled by only a few foreign suppliers, Washington's ability to impose sanctions that undermine the policies of European countries,[39] and the increasing tendency of China to deploy economic coercion to assert its national interests.[40] Many Europeans are also made uneasy by the potential for machine learning to exacerbate discrimination in society.[41] This is because data may reflect existing structural inequalities and the realities of an unjust society, which can become embedded in AI systems and exacerbate discriminatory outcomes.[42] Another concern is that inadequate data protections deployed by businesses outside the EU have allowed countries like Russia to exploit them for political purposes to target disinformation campaigns and influence elections in Europe.[43] Reconciling these concerns with the benefits of data access and re-use for machine learning will be an ongoing challenge for the EU.[44]

There is the risk that if the EU fails to resolve these tensions through data rules that foster innovation, businesses could face higher costs to deploy machine learning, weakening their ability to compete in a global market. The resulting lack of European innovation could reinforce the EU's dependencies on foreign digital technologies, with local industries taking a smaller slice of the global value chain. However, there is also a risk that the EU implements short-term, innovation-driven policies that could embed economic systems that fail to protect fundamental European rights, exacerbate social inequalities, and threaten democratic institutions and processes.[45] If the EU cannot get this balance right, its inability to present an effective regulatory model will undermine its ability to shape global rules at the intersection of technology and society.

Aware of these costs, the European Commission has made "a Europe fit for the digital age" a priority to achieve by 2024.[46] The EU has proposed a raft of legislative proposals, with more in the pipeline, informed by the commission's *European Strategy for Data* and *White Paper on Artificial Intelligence*, both released in 2020, including the Data Governance Act proposal introduced in November 2020,[47] and the Artificial Intelligence Act proposal introduced in April

35 Tim Hwang, *Shaping the Terrain of AI Competition,* Center for Security and Emerging Technology (CSET), Georgetown University's Walsh School of Foreign Service, June 2020, 17–18, https://cset.georgetown.edu/research/shaping-the-terrain-of-ai-competition/.

36 Data markets are where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

37 "The European Data Market Study Update," European Commission (website).

38 Long Zhang, "China's Big Data Market to Continue Expansion," Xinhua News Service, March 14, 2021, https://www.shine.cn/biz/economy/2103145919/.

39 Ellie Geranmayeh and Manuel Lafont Rapnouil, "Meeting the Challenge of Secondary Sanctions," Policy Brief, European Council on Foreign Relations, June 25, 2019, https://ecfr.eu/publication/meeting_the_challenge_of_secondary_sanctions.

40 James Laurenceson, "Will the Five Eyes Stare Down China's Economic Coercion?," Lowy Institute, accessed July 8, 2021, https://www.lowyinstitute.org/the-interpreter/will-five-eyes-stare-down-china-s-economic-coercion.

41 Kantar, "Standard Eurobarometer 92: Report on Europeans and Artificial Intelligence," Kantar (consulting company), November 2019, 19, https://europa.eu/eurobarometer/api/deliverable/download/file?deliverableId=72485.

42 Genevieve Smith and Ishita Rustagi, *Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook*, Berkeley Haas Center for Equity, Gender and Leadership, July 2020, 23, https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf.

43 Francesca Bignami, "Schrems II: The Right to Privacy and the New Illiberalism," *Verfassungsblog*, accessed April 20, 2021, https://verfassungsblog.de/schrems-ii-the-right-to-privacy-and-the-new-illiberalism/.

44 Hwang, *Shaping the Terrain*, 16–18.

45 Hwang, *Shaping the Terrain*, 13–15.

46 European Commission, "6 Commission Priorities for 2019-24," European Commission website, https://ec.europa.eu/info/strategy/priorities-2019-2024_en.

47 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act), November 25, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=EN.

2021.[48] It has also allocated €143.4 billion ($166.25 billion) toward the EU "single market, innovation, and digital" between 2021 and 2027.[49] In pursuing this agenda, the commission needs to balance stimulating innovation with mitigating the risks machine learning can pose to fundamental rights, European autonomy, and society. This report proposes a path forward for the EU against the backdrop of rapid European rulemaking and intensifying competition between China and the United States in data and machine learning.

Our approach examines the EU's geopolitical considerations, domestic considerations of EU member states including societal interests and the protection of individual rights, and the commercial considerations of EU-based businesses and makes recommendations designed to navigate the complex equities at stake. While our research takes a broad view, we sought specific examples, concerning autonomous vehicles and healthcare, to demonstrate the benefits and risks to society posed by the deployment of machine learning. We similarly sought examples from four countries—Finland, France, Germany, and Poland—to highlight the diversity of priorities, concerns, and approaches among EU member states.

The report proposes three broad objectives for the EU to pursue in developing data rules for machine learning:

1) leverage the scale of Europe
2) balance economic autonomy and openness
3) protect fundamental rights

Under these categories, we recommend fifteen specific actions for the European Commission.

It is important to note that framing the issue of data governance around EU competition with the United States and China can lend itself to lines of argument that downplay the importance of protecting the rights of individuals to control their own data and the social implications of data-driven decision making based on machine learning systems. To make the analysis and recommendations more easily digestible, they are separated into seven chapters under three themes; however, all the sections are strongly interconnected and are not designed to be read in isolation from one another. We have attempted to flag in the analysis the cases in which lines of reasoning risk downplaying important social interests, which can be complex and multidimensional, and have linked them to other parts of the report that discuss those issues in more detail.

48 European Commission sources: Artificial Intelligence Act, April 21, 2021, https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF; *European Strategy for Data*, February 19, 2020, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf; and *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*, 2020, https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

49 European Commission, "EU's Next Long-Term Budget & NextGenerationEU: Key Facts and Figures," November 11, 2020, accessed July 31, 2021, https://ec.europa.eu/info/sites/default/files/about_the_european_commission/eu_budget/mff_factsheet_agreement_en_web_20.11.pdf.

Part One: Leverage the Scale of Europe

1.1 Increase access to public-sector data

Public-sector data has significant potential for commercial and research purposes including to develop new products and services with machine learning.[50] The scale of the EU single market offers opportunities to pool data resources, the reuse of which could contribute up to €194 billion to the EU economy by 2030.[51] However, **the EU needs to address the complexity of navigating fragmented open data rules and procedures across Europe and, with the appropriate privacy-preserving techniques and protections in place, enable limited access to sensitive data,[52] if it is to innovate on par with China and the United States.**

To achieve this, the European Commission should establish an EU data service that facilitates access to publicly held data, helping to reduce the administrative and legal complexity for businesses seeking to reuse nonsensitive public-sector data. The service could also work with national data-protection authorities and public bodies to make sensitive data more widely available for research and development purposes under adequate security and privacy-preserving protections that meet data-protection requirements, such as the GDPR.

This chapter should be read in conjunction with chapter 2.3 on data ecosystems, which looks at the implications of data sharing among businesses, as well as chapter 3.1 on privacy and trust, which has a greater focus on the risks to individuals, as both chapters are relevant to the issues discussed in this chapter.

EU Policy Context

There are fragmented rules and processes on accessing and using public-sector data in Europe. The EU's 2019 *Directive on Open Data and the Re-Use of Public Sector Information* updated the minimum standard for public bodies to make nonsensitive data accessible to the public.[53]

However, while the directive strengthens language against imposing conditions on the reuse of data unless justified on public-interest grounds, it does not require the use of standard licenses that make it easier for organizations to use and combine data sets.[54] The directive also implies that those seeking EU-wide data sets need to approach each public-sector body individually. Furthermore, it does not clarify conditions to access to sensitive data (see figure 2 for examples of sensitive data). The commission's proposed Data Governance Act seeks to address this gap by clarifying obligations for the public sector to share sensitive data that falls outside of the scope of the Open Data Directive.[55] The commission has also foreshadowed sectoral legislation to create "common data spaces" to make public and private data more widely available.[56] It envisages these common data spaces will provide nondiscriminatory access to high-quality data between businesses and with government.[57]

Geopolitical Considerations

Barriers to accessing public-sector data put Europe at a disadvantage vis-à-vis China and the United States. Charles Michel, president of the European Council, has described the current unavailability and fragmentation of open data policies as the "weak point" of the European market.[58] Compared to the EU, Chinese organizations benefit from the ability to extract vast amounts of data from users owing to few and lax privacy laws.[59] The US private sector has amassed troves of data drawn from its dominant position in most sectors in the United States and globally.[60] For example, in a conversation with experts Marc Lange

50 "Open Data," Shaping Europe's Digital Future (webpages), European Commission, March 10, 2021, https://digital-strategy.ec.europa.eu/en/policies/open-data-0.

51 "From the Public Sector Information (PSI) Directive to the Open Data Directive," Shaping Europe's Digital Future (webpages), European Commission, March 10, 2021, https://digital-strategy.ec.europa.eu/en/policies/psi-open-data.

52 Sensitive data can include personal data described under the General Data Protection Regulation, as well as intellectual property, and information that is commercially confidential, statistically confidential, or relates to national security under the Open Data and the Reuse of Public Sector Information Directive.

53 Directive on Open Data and the Re-Use of Public Sector Information, *Official Journal of the European Union* 62, no. L172 (June 26, 2019): 56–83, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2019:172:FULL&from=EN.

54 Directive on Open Data and the Re-Use of Public Sector Information.

55 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act).

56 European Commission, "A European Strategy for Data," Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions COM(2020), 66, February 19, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN.

57 European Commission, Proposal for a Regulation Laying down Harmonised Rules, 29.

58 Charles Michel, keynote address, *Masters of Digital 2021*, Digital Europe (trade association summit), February 3, 2021, https://www.youtube.com/watch?v=TR0p_9Ii0U8.

59 Hwang, *Shaping the Terrain*.

60 Matt Sheehan, "Much Ado About Data: How America and China Stack Up," MacroPolo, July 16, 2019, https://macropolo.org/ai-data-us-china/.

Figure 2: Examples of Sensitive Data

Data	Examples
Under the General Data Protection Regulation (GDPR)[1]	
Personal Information	Name, gender, photo ID, and other economic, cultural, and social identifiers, and behaviors.
Under the EU Open Data and the Re-use of Public Sector Information Directive[2]	
Classified Information	Information related to national security, defense, or public security.
Intellectual Property and Commercially Confidential Information	Engineering designs, service delivery data, city planning application, and tender submissions.
Statistical Confidentially	When matching with another database could reveal the identity or sensitive attributes of a person or entity.

1 General Data Protection Regulation, Official Journal of the European Union, Chap. 1, Article 4, accessed March 21, 2021, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2018:0232:FIN#footnote43.
2 Directive on Open Data and the Re-Use of Public Sector Information," Chap. 1, Article 1.

and Luc Nicolas from the European Health Telematics Association (EHTEL), they noted that US healthcare companies often had direct access to data given the heterogeneity and market orientation of the different health systems, giving US companies greater capacity to directly deploy machine learning in healthcare. They observed that health data in Europe was generally only open to each hospital's research department, which, given their nonprofit legal status, sought mainly to publish scientific research rather than develop healthcare products and services.[61] Given these systemic differences, to compete with China and the United States, the EU needs to better enable companies to train machine learning algorithms with data held by the public sector.

National Considerations

EU member states have different levels of comfort with allowing access to sensitive data. Member states are concerned about ensuring and guaranteeing the trust and confidence of citizens and businesses (see chapter 3.1). For instance, countries where citizens have strong concerns about privacy, like Germany, have until recently been reluctant to enable public access to data in sensitive industries such as healthcare.[62]

On the other hand, countries like Finland and Estonia lead the way in digital health, and health providers already exchange some patient data.[63] In 2019, the Finnish Act on the Secondary Use of Health and Welfare Data entered into force to facilitate access to social and health data, the first such legislation by an EU member state.[64] The act establishes a central data-permit authority for the secondary use of health and welfare data, unifying access procedures and decision making.[65] Given these different levels of comfort, the European Commission will need to balance increased data availability while ensuring data-protection measures satisfy the higher expectations of some member states.

Business Considerations

Low-cost access to multiple public-sector databases can enable businesses to develop new products and services. Reaching the scale and scope needed to train

61 Marc Lange and Luc Nicolas, Interview with the European Health Telematics Association (EHTEL), February 12, 2021.
62 Carmen Paun, "Germany Aims to Become EU Leader in Digital Health Revamp," *Politico*, October 16, 2019.
63 Paun, "Germany Aims to Become EU Leader."
64 "Secondary Use of Health and Social Data," Finland's Ministry of Social Affairs and Health (website), April 16, 2021, https://stm.fi/en/secondary-use-of-health-and-social-data.
65 Saara Malkamäki and Hannu Hämäläinen, "New Legislation Will Speed up the Use of Finnish Health Data," Sitra, May 14, 2019, https://www.sitra.fi/en/blogs/new-legislation-will-speed-use-finnish-health-data/.

an algorithm may require pooling data from a variety of open data sources. However, EU member states each approach the openness of public data differently,[66] meaning businesses need to treat data under different conditions, increasing legal complexity and operational costs.[67] One example is the range of different open data licenses used, which is estimated to reach ninety different licenses used across various national, regional, and municipal governments in the EU.[68] If businesses can more simply access and use open data across the EU, it would make it easier for businesses to combine data sets to train new products and services based on machine learning.

Businesses would benefit from simpler procedures to access sensitive data under conditions that maintain data protection. Implementing simple procedures for the public to access sensitive data under the right conditions can enable the development of innovative machine learning products and services. For example, Corti is a natural language processing start-up that learns from medical conversations in emergency calls to better and sooner detect if someone is having a heart attack by listening to key words, background noises, and emotions.[69] Copenhagen Emergency Health Services shared audio data from hundreds of thousands of emergency calls with Corti to enable machine learning; in trials, the software reduced the number of undetected out-of-hospital heart attacks in Copenhagen by 43 percent.[70] Corti was further piloted in Italy and France, and while meeting GDPR posed significant challenges, it was able to work with emergency services to meet European data-protection requirements.[71] This demonstrates that enabling access to sensitive data with the right protections for privacy (see chapter 3.1) can lead to innovations that benefit consumers and society.

There are significant risks in allowing access to sensitive data, but privacy-preserving machine learning techniques are emerging. Once data is released, controlling who can view data and for what purpose becomes extremely difficult to enforce, opening opportunities for misuse.[72] Data scientists are developing privacy-preserving techniques that can still produce accurate machine learning outputs, ranging from anonymization that removes sensitive entries and pseudonymization that replaces them with made-up ones—though these have been proven to be vulnerable to reidentification[73]—to more sophisticated decentralized, differential, or encryption techniques.[74] Applying these privacy-preserving techniques as a condition to access to sensitive public data could provide opportunities for researchers to develop and deploy beneficial machine learning applications while protecting sensitive data.

Recommendations for the European Commission

A. Establish an EU data service under legislation creating common data spaces to facilitate access to public-sector data in consultation with national data -protection authorities, starting with the health data space planned for the end of 2022.[75]

Piloting the new system in the healthcare sector would enable the authority to draw lessons from Finland's permit-granting authority (Findata), before expanding its scope to cover other sectors. The service could be an expansion of the EU's open data portal (data.europe.eu) and be jointly overseen by the European Data Protection Board, as an independent European body focused on the consistent application of data-protection rules throughout the EU,[76] and the European Data

66 "Directive on Open Data and the Re-Use of Public Sector Information., Paragraph 15.

67 "Report of License Proliferation Committee and Draft FAQ," Open Source Initiative (website), accessed April 17, 2021, https://opensource.org/proliferation-report.

68 Rufus Pollock and Danny Lämmerhirt, "Open Data Around the World: European Union," State of Open Data (research project funded by the International Research Centre with the support of the Open Data for Development [OD4D] Network), accessed April 17, 2021, https://www.stateofopendata.od4d.net/chapters/regions/european-union.html.

69 Bernard Marr, "AI That Saves Lives: The Chatbot That Can Detect A Heart Attack Using Machine Learning," accessed April 17, 2021, https://bernardmarr.com/default.asp?contentID=1762.

70 "Corti Website: About," accessed April 17, 2021, https://www.corti.ai/about.

71 European Emergency Number Association (EENA), "Detecting Out-of-Hospital Cardiac Arrest Using Artificial Intelligence Protect Report," January 23, 2020, 14–15, https://eena.org/document/detecting-out-of-hospital-cardiac-arrest-using-artificial-intelligence/.

72 Ari Ezra Waldman, "Privacy Law's False Promise," *Washington University Law Review* 97, no. 3, (2020), 773–834, https://openscholarship.wustl.edu/law_lawreview/vol97/iss3/7.

73 Caroline Perry, SEAS Communications, "You're Not So Anonymous," *Harvard Gazette*, October 18, 2011, https://news.harvard.edu/gazette/story/2011/10/youre-not-so-anonymous/.

74 For example, federated machine learning is a decentralized approach that distributes the algorithm to where the data is, rather than gathering the data where the algorithm is, removing the need to transfer the data. Another example is differential privacy, which randomly shuffles data to remove the association between individuals and their data entries in a way that specifically designed algorithms can still conduct statistical analysis on the whole data set. Homomorphic encryption, which allows computation on encrypted data as if it were unencrypted is another approach, albeit technically challenging. See Georgios A. Kaissis et al., "Secure, Privacy-Preserving and Federated Machine Learning in Medical Imaging," *Nature Machine Intelligence* 2, no. 6 (June 2020): 305–11, https://doi.org/10.1038/s42256-020-0186-1.

75 European Commission, "European Health Data Space," European Commission (website), accessed April 17, 2021, https://ec.europa.eu/health/ehealth/dataspace_en.

76 "Who We Are," European Data Protection Board (website), accessed August 2, 2021, https://edpb.europa.eu/about-edpb/about-edpb/who-we-are_en.

Innovation Board in the proposed Data Governance Act, given its intended focus on data standardization.[77] The authority's functions could include:

- Serving as a single window for data-access requests for data sets across multiple jurisdictions and coordinate access to various data sets under single or at least compatible format and license
- Assessing requests for data not available for open access, including sensitive data, in consultation with national data-protection authorities and make recommendations to relevant public bodies on whether to allow access and under what conditions, such as the use of specific privacy-preserving techniques and security precautions

The service could operate on a cost-recovery basis, passing expenses on to the businesses and research institutes using the service, which would still help reduce costs arising from administrative and legal complexity for businesses seeking to reuse public-sector data. It would also increase the availability of sensitive data for research and development while ensuring adequate protections laid out in relevant national-level data-protection legislation, for example, through privacy-preserving machine learning techniques.

1.2 Harmonize data standards for interoperability

Machine learning and the Internet of Things (IoT) will bring artificial intelligence to our devices, factories, hospitals, and homes, creating efficiencies and driving productivity.[78] To achieve this, however, the EU needs to enable the sharing of high-quality data between distributed devices and systems.[79] **The lack of harmonized data standards across the EU[80] slows the deployment of machine learning compared to China and the United States.[81]**

To address this, the European Commission should facilitate the development of sector-specific data models, application programming interfaces (APIs) and open source software that enable the seamless exchange and translation of differently structured data across systems and organizations.

EU Policy Context[82]

The European Commission's efforts to encourage the voluntary adoption of data interoperability standards have made slow progress. For example, decadelong efforts by the commission under the European Patient Smart Open Services (epSOS) project to encourage the implementation of interoperable health data standards have been stymied by reluctant healthcare IT providers and ambivalent national governments.[83] More recently, the commission enacted the 2019 Recommendation on a Europe Electronic Health Record exchange, but it lacks legal enforcement to incentivize compliance.[84]

Looking forward, the EU's 2020 Data Strategy included the development of rules for common "data spaces" intended to "set up structures that enable organizations to share data." The European Commission's 2020 proposal for a Data Governance Act also includes the establishment of a European Data Innovation Board that would focus on data standardization.[85] However, it will be very difficult for member states to agree to a standard data format due to constraints by existing IT infrastructure as well as incompatible and inflexible national privacy and security regulations.[86] **To facilitate widespread adoption of machine learning, the commission needs to find a way to enable data interoperability while avoiding**

77 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act).

78 "Generating Value at Scale with Industrial IoT," McKinsey Digital (website), McKinsey & Company, February 5, 2021, https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/a-manufacturers-guide-to-generating-value-at-scale-with-industrial-iot.

79 Gudivada, Apon, and Ding, "Data Quality Considerations for Big Data."

80 Daniel Rubinfeld, "Data Portability," *Competition Policy International*, November 26, 2020, https://www.competitionpolicyinternational.com/data-portability/.

81 Winston Maxwell et al., *A Comparison of IoT Regulatory Uncertainty in the EU, China, and the United States*, Hogan Lovells (law firm including Hogan Lovells International LLP, Hogan Lovells US LLP, and affiliated businesses), March 2019.

82 This chapter focuses on electronic health records (EHR) as a case study on data interoperability. The siloed nature of healthcare IT systems, combined with incompatible national privacy and security rules, provide a good demonstration of the challenges the EU will face more broadly in connecting distributed data systems to leverage the benefits of machine learning.

83 "EpSOS," healthcare-in-europe.com (platform website), accessed 2021, https://healthcare-in-europe.com/en/news/epsos.html.

84 European Commission, Recommendation on a European Electronic Health Record Exchange Format, C(2019)800 of February 6, 2019, European Commission (website), accessed 2021, https://digital-strategy.ec.europa.eu/en/library/recommendation-european-electronic-health-record-exchange-format.

85 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act).

86 Philipp Grätzel von Grätz, "Transforming Healthcare Systems the European Way," *Healthcare IT News*, May 24, 2019, https://www.healthcareitnews.com/news/emea/transforming-healthcare-systems-european-way.

costly and complex overhauls of member state regulations and IT systems.

Geopolitical Considerations

China uses different digital standards than the EU and is internationalizing those standards through global bodies and bilateral agreements, making its digital exports more interoperable with digital systems in foreign markets. The EU has a strong voice on international data standards through international bodies like the International Standards Organization (ISO) and the International Electrotechnical Committee (IEC).[87] However, China is rapidly increasing its participation in standard setting and using other strategies to influence standards in export markets.[88] Björn Fägersten and Tim Rühlig at the Swedish Institute of International Affairs outline how China is rapidly developing industry standards and proposing new work items in ISO and IEC based on nascent technologies to gain first mover advantage in technical standardization.[89] China also is exporting its domestic standards through foreign investments such as Belt and Road Infrastructure projects, including incorporating clauses on standardization in government-to-government memorandums of understanding.[90] Given the scale of the EU, standards adopted EU-wide can create market power that sway smaller economies to adopt EU standards.[91] However, if data standards remain fragmented in Europe, they are less appealing to other countries compared to Chinese or other technical standards that are already widespread.

National Considerations

Capacity and existing regulations constrain some EU member states from implementing common data standards. Each EU member state has its own approach to public data infrastructure that uses disparate data formats based on legacy systems, as well as national security standards tied to privacy requirements,[92] to which amendments can be controversial.[93] For example, Germany's mandatory technical specifications for electronic health records (EHR) are different from those encouraged by the European Commission due to security requirements.[94] For other member states, regulatory constraints may be surmountable but there is a lack of organizational capacity to adopt interoperability standards. For example, while Poland faces fewer legal barriers,[95] its hospitals and clinics have been slower to digitize health data,[96] resulting in a Polish law requiring national adoption of EHR by 2014 to be delayed by three years.[97] In general, studies have found that Central and Eastern European countries continue to face difficulties operationalizing EHR due to fragmented policy frameworks and constraints on financial resources.[98]

Business Considerations

Reducing the time and costs involved in preparing data for machine learning will make the technology more accessible. Preparing large data sets for machine learning can be difficult, resource intensive, and costly for business. Currently, "data wrangling" takes up about 80 percent of the time consumed in most artificial intelligence projects, as data from various data sources may need to be manually labeled and structured in a coherent way.[99] Data also needs semantic interoperability, whereby the labels or categories mean the same thing across databases.[100] There are solutions to this problem in the form of data models like the Observational Medical Outcomes Partnership (OMOP) that can translate different data formats so long as the semantic vocabulary is the same (e.g., using the term "myocardial

87 Björn Fägersten and Tim Rühlig, *China's Standard Power and Its Geopolitical Implications for Europe,* Swedish Institute of International Affairs, 2019.

88 Fägersten and Rühlig, *China's Standard Power*.

89 Fägersten and Rühlig, *China's Standard Power*.

90 Fägersten and Rühlig, *China's Standard Power*.

91 "The EU Wants to Set the Rules for the World of Technology," *Economist*, February 20, 2020, http://www.economist.com/business/2020/02/20/the-eu-wants-to-set-the-rules-for-the-world-of-technology.

92 Philipp Grätzel von Grätz, "Transforming Healthcare Systems the European Way."

93 Grätzel von Grätz, "Transforming Healthcare Systems."

94 Grätzel von Grätz, "Transforming Healthcare Systems."

95 Milieu Ltd. and Time.lex, *Overview of the National Laws on Electronic Health Records in the EU Member States: National Report for Poland*," Prepared for the Executive Agency for Health and Consumers, March 5, 2014, https://ec.europa.eu/health/sites/health/files/ehealth/docs/laws_poland_en.pdf.

96 Leontina Postelnicu, "EU Report Looks at Uptake of EHRs, EPrescribing, and Online Access to Health Information," *Healthcare IT News*, December 14, 2018, https://www.healthcareitnews.com/news/emea/eu-report-looks-uptake-ehrs-eprescribing-and-online-access-health-information.

97 Aleksandra Czerw et al., "Implementation of Electronic Health Records in Polish Outpatient Health Care Clinics: Starting Point, Progress, Problems, and Forecasts," *Annals of Agricultural and Environmental Medicine* 23, no. 2 (2016): 6.

98 Marek Ćwiklicki et al., "Antecedents of Use of E-Health Services in Central Eastern Europe: A Qualitative Comparative Analysis," *BMC Health Services Research* 20, no. 1 (March 4, 2020): 171, https://doi.org/10.1186/s12913-020-5034-9.

99 Tom Gauld, "For AI, Data Are Harder to Come by Than You Think," *Economist*, June 13, 2020, http://www.economist.com/technology-quarterly/2020/06/11/for-ai-data-are-harder-to-come-by-than-you-think.

100 "What Does Interoperability Mean for the Future of Machine Learning?," Appen (blog on corporate website), September 22, 2020, https://appen.com/blog/what-does-interoperability-mean-for-the-future-of-machine-learning/.

infarction" rather than "heart attack").[101] These kinds of data models have the potential to enable smoother compilation of large-scale data sets. In doing so, they make large amounts of data more easily available to machine leaning algorithms resulting in accurate outputs, such as a higher likelihood of correctly diagnosing a disease.

Interoperability—the ability for two systems to communicate effectively—is key for machine learning.[102] For instance, patients are increasingly tracking and generating large volumes of personal health data through wearable sensing and mobile health (mHealth) apps.[103] One example is KardiaMobile, a personal ECG device that can monitor heart rhythms and instantly detect irregularities, such as atrial fibrillation that can lead to a stroke.[104] The Kardia app allows patients to track ECG data over time and share the recordings directly with their physician.[105] However, a European study of health clinicians using mHealth apps with their patients found that they did not have the means to transfer this data to the patient's electronic health record.[106] This means a patient's data from a device cannot be compiled with the patient's other health data for further research and development purposes, which weakens the potential for machine learning to improve patient health outcomes.

Open source data-management software—which are characterized by open licenses, commitment to standards, and collaboration among developers—could offer a pathway to help institutions keep pace with technological advancements and enhance interoperability.[107] Compared to proprietary software, the open architecture of open source software can more easily allow for the integration of various products like medical devices and wearables into a single system, and commercial versions can adapt the software to support the needs of specific users.[108] Open source software can be vulnerable to hacking and data breaches when not maintained by an active community;[109] however, open source projects with an active user base are generally more likely to discover bugs and security vulnerabilities, and fix them more quickly, than most proprietary software (i.e., closed source software).[110]

Recommendations for the European Commission

A. Mandate and fund the European Data Innovation Board (in the proposed Data Governance Act) to establish an EU technical data standards registry to facilitate the sharing of various technical data standards between national standards organizations, industry bodies, associations, and businesses.

The purpose of the registry would be to increase transparency of the various data standards at the European, national, sectoral, association, and enterprise levels that are in use across the EU. This would assist technical experts identify opportunities and develop ways to enhance interoperability. The registry could draw on the example of the registry maintained by the Standardization Administration of China that includes government, industry regulator, local, association, and enterprise-level standards.[111] Another example is the comprehensive list of standards developed by the Canadian Data Governance Standardization Collaborative.[112]

B. Mandate and fund the European Data Innovation Board (proposed under the Data Governance Act) to establish sector-specific expert committees to recommend data interoperability standards, drawing on technical experts from standardization bodies, industry, academia, and consumer groups.

The expert committees could be established alongside each "common data space" and should specifically focus on:

101 Observational Health Data Sciences and Informatics, "OMOP Common Data Model," accessed April 18, 2021, https://www.ohdsi.org/data-standardization/the-common-data-model/.

102 "What Does Interoperability Mean for the Future of Machine Learning?," Appen blog.

103 Haining Zhu et al., "Sharing Patient-Generated Data in Clinical Practices: An Interview Study," *AMIA Symposium* 2016 (2016): 1303–12.

104 Sushravya Raghunath et al., "Deep Neural Networks Can Predict Incident Atrial Fibrillation from the 12-Lead Electrocardiogram and May Help Prevent Associated Strokes," Preprint, submitted April 27, 2020, https://doi.org/10.1101/2020.04.23.20067967.

105 EIT Health and McKinsey & Company, *Transforming Healthcare with AI*, March 2020, https://thinktank.eithealth.eu/wp-content/uploads/2020/12/EIT-Health-and-McKinsey-%E2%80%93-Transforming-Healthcare-with-AI.pdf.

106 Zhu et al., "Sharing Patient-Generated Data in Clinical Practices."

107 Maria Sukhova, "Pros and Cons of Open Source Software in Healthcare," Blog, Auriga (outsourcing software company), April 19, 2017, accessed August 6, 2021, https://auriga.com/blog/2017/pros-and-cons-of-open-source-software-in-healthcare/.

108 Sukova, "Pros and Cons of Open Source Software in Healthcare."

109 "Open Source Does Not Equal Secure: Schneier on Security," Blog, accessed August 13, 2021, https://www.schneier.com/blog/archives/2020/12/open-source-does-not-equal-secure.html.

110 Ron Rymon, "Why Open Source Software Is More Secure Than Commercial Software," WhiteSource (blog), April 7, 2021, accessed August 13, 2021, https://www.whitesourcesoftware.com/resources/blog/3-reasons-why-open-source-is-safer-than-commercial-software/.

111 The US-China Business Council, *Standards Setting in China: Challenges and Best Practices*, 2020, 22.

112 Standards Council of Canada, *Canadian Data Governance Standardization Roadmap*, Annex B, 75–155, 2021, https://www.scc.ca/en/system/files/publications/SCC_Data_Gov_Roadmap_EN.pdf.

i. **Standard data models** that enable the translation of differently structured data across systems and organizations. An example of such a data model is the OMOP model for health data mentioned above.

ii. **Standard application programming interfaces (APIs)** that enable apps and databases to seamlessly exchange data. For example, the United States requires health information to be made available by the Fast Healthcare Interoperability Resources (FHIR) API standard.[113]

iii. **Free open source data-management software** that embed sector-specific standards, encourage a community of practice, and allow for commercial adaption to serve the needs of specific users.

The widespread adoption of common data models would enable EU member states to implement data interoperability standards without overhauling costly IT infrastructure and amending sensitive privacy and security regulations. Combined with common APIs and free open source data-management software this would reduce costs and increase accessibility to data for businesses to adopt machine learning. The broad adoption of these standards in the EU would also increase the chances of European data interoperability standards being adopted internationally.

113 United States Department of Health and Human Services, Centers for Medicare & Medicaid Services, "Interoperability and Patient Access Final Rule," 85 Fed. Reg. 25510 (May 1, 2020), https://www.federalregister.gov/documents/2020/05/01/2020-05050/medicare-and-medicaid-programs-patient-protection-and-affordable-care-act-interoperability-and.

Part Two: Balance Economic Autonomy and Openness

2.1 Diversify data storage and processing

Cloud service providers increasingly become geopolitical actors whose influence shapes economic growth, international security competition, and global balance of power.[114] To ensure redundancy in essential data services and reduce vulnerability to potential coercion, the European Union needs to diversify its dependency on a few non-European cloud service providers. However, in doing so, **the EU needs to maintain competitive cloud offerings for European businesses and avoid succumbing to calls for protectionist measures that potentially harm businesses and could negatively impact some EU member states.**

The European Commission should consider further enhancing investment into the existing federated cloud project to diversify the supply of data infrastructure services. It also needs to ensure coordination of proposed and existing initiatives in this area, such as Gaia-X, and aim to minimize the compliance costs of new regulatory measures that could disadvantage European cloud providers. It is worth noting that within the EU policy debate there is a strong focus on protecting data from foreign jurisdictions and weaker data-protection regimes, often mentioned in the context of data localization efforts that force data to be hosted in the jurisdiction it was collected.[115] Therefore, this chapter should be read in conjunction with chapters 2.2 on simplifying data transfers and with chapter 3.1 on enhancing privacy and trust.

EU Policy Context

The European Commission, supported by most EU member states, aims to reduce dependencies on external cloud providers by setting new regulations and funding "federated cloud infrastructure."[116] Building on the data strategy proposed by the commission, twenty-five EU member states signed a joint declaration in October 2020 pledging to invest up to €10 billion in the cloud sector and establish the "European Alliance on Industrial Data, Edge and Cloud."[117] To shape these efforts, twenty-seven CEOs of European companies shared priority areas for cloud and edge ecosystem development and investment with European Commissioner for Internal Market Thierry Breton in May 2021.[118]

In July 2021, the commission officially launched the European Alliance on Industrial Data, Edge and Cloud, which aims to combine resources from existing EU programs, industry and member states, for the creation of a European Federated Cloud in 2021.[119] The alliance plans to write a Cloud rulebook that would provide a single European framework of cloud use and interoperability requirements, and it has posed the idea of a European marketplace for cloud offerings: a single portal to cloud services meeting key EU standards and rules.[120] There also is a strong focus on so-called edge computing, which instead of relying on centralized infrastructure, allows for real-time data sharing and analysis directly between devices. In addition to this, the 2020 Digital Markets Act proposed by the commission aims to mitigate the risk of dependencies on cloud providers through stricter portability and data access obligations.[121]

It is yet unclear whether the European Alliance on Industrial Data, Edge and Cloud will be able to coordinate the efforts needed to compete with the leading

114 Trey Herr, *Four Myths about the Cloud: The Geopolitics of Cloud Computing*, Atlantic Council, August 31, 2020, https://www.atlanticcouncil.org/in-depth-research-reports/report/four-myths-about-the-cloud:-the-geopolitics-of-cloud-computing/.

115 Herr, *Four Myths about the Cloud*.

116 Burwell and Propp, *The European Union and the Search for Digital Sovereignty*.

117 "Commission Welcomes Member States' Declaration on EU Cloud Federation," Press Release, European Commission (website), October 15, 2020, https://digital-strategy.ec.europa.eu/en/news/commission-welcomes-member-states-declaration-eu-cloud-federation; and Melissa Heikkila and Janosch Delcker, "EU Shoots for €10B 'Industrial Cloud' to Rival US," *Politico*, October 15, 2020, https://www.politico.eu/article/eu-pledges-e10-billion-to-power-up-industrial-cloud-sector/.

118 European Commission, "Today the Commission Receives Industry Technology Roadmap on Cloud and Edge," Report Announcement, Shaping Europe's Digital Future (webpages), May 2021, https://digital-strategy.ec.europa.eu/en/library/today-commission-receives-industry-technology-roadmap-cloud-and-edge.

119 Heikkila and Delcker, "EU Shoots for €10B 'Industrial Cloud' to Rival US."

120 "Cloud Computing," Shaping Europe's Digital Future (webpages), European Commission, March 10, 2021, https://digital-strategy.ec.europa.eu/en/policies/cloud-computing.

121 European Commission, Directorate-General for Communications Networks, Content, and Technology (DG CNECT), Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act), 2020/0374/COD (2020), https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1608116887159&uri=COM%3A2020%3A842%3AFIN.

non-European cloud providers. There is also a question of how specifically the Alliance relates to other initiatives, such as the Franco-German Gaia-X cloud network. Gaia-X aims to serve as a platform and Europe-wide network joining up cloud services providers, high performance computing, and edge system providers from a number of EU-based and international organizations.[122] It is crucial that these policy and funding initiatives are coherent and that potential overlaps are minimized to enable data flows while maintaining adequate data protections.[123]

Geopolitical Considerations

The concentration of the EU's cloud services among few US suppliers makes it vulnerable to external supply disruptions. According to Atlantic Council research, some 92 percent of the Western world's data are stored in the United States.[124] Just three US providers—Amazon, Microsoft, and Google—account for 66 percent of the European cloud market.[125] EU-based cloud providers have seen their share of the European market decline from 26 percent in 2017 to 16 percent in 2020 (see figure 3) and the largest EU-based provider—Deutsche Telekom— accounts for only 2 percent of the market.[126]

At the same time, the uptake of cloud among EU businesses has been on the rise, which highlights the urgency of addressing the issue.[127] As cloud services become part of the critical infrastructure, they become increasingly vulnerable to disruption.[128] For example, a single cloud provider's supply chain decisions and internal security policies can impact millions of customers.[129] As the United States has previously threatened to cut European companies off from critical US services for engaging in business activities condoned by European countries,[130] the growing dependence on a few US providers is increasingly seen as a strategic weakness that could undermine European foreign policies and national interests.[131]

Europeans are also concerned about foreign authorities potentially accessing their data and risks to data protection. The emergence of foreign entities holding large troves of data on individuals is a growing concern for Europeans who were alarmed by Edward Snowden's disclosure of electronic mass surveillance by the US government.[132] These concerns have intensified since the adoption of the US Clarifying Lawful Overseas Use of Data Act (CLOUD Act) that can require US tech companies to hand data to US authorities, even if it is stored outside the United States.[133] For example, in April 2019, Germany's top data protection office told *Politico* that Amazon's cloud hosting services were not suitable for storing German police data due to the risks that US authorities would access the data.[134] In May 2021, the EU privacy watchdog, the European Data Protection Supervisor, launched an investigation looking into compliance of EU bodies' using of cloud computing services provided by Amazon and Microsoft due to concerns about the transfer of personal data to the United States.[135] In addition to this, the EU is concerned about cloud service providers being subject to legislation that would contradict the EU's General Data Protection Regulation and therefore weaken existing data protection enforcement, in particular in the context of concerns about several Chinese cybersecurity and national intelligence laws (see chapter 2.2).[136]

122 "GAIA-X: A Federated Data Infrastructure for Europe," accessed August 10, 2021, https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html; and Janosch Delcker and Melissa Heikkila, "Germany, France Launch Gaia-X Platform in Bid for 'Tech Sovereignty,' " *Politico*, June 4, 2020, https://www.politico.eu/article/germany-france-gaia-x-cloud-platform-eu-tech-sovereignty/.

123 Andrea Renda et al., *The Digital Transition: Towards a Resilient and Sustainable Post-Pandemic Recovery,* Centre for European Policy Studies (CEPS), 2021, https://www.ceps.eu/wp-content/uploads/2021/07/The-Digital-Transition_CEPS-TF-WGR.pdf.

124 Burwell and Propp, *The European Union and the Search for Digital Sovereignty.*

125 "European Cloud Providers Struggle to Reverse Market Share Losses," Synergy Research Group (website), January 14, 2021, https://www.srgresearch.com/articles/european-cloud-providers-struggle-reverse-market-share-losses.

126 "European Cloud Providers Struggle," Synergy.

127 "Cloud Computing: Statistics on the Use by Enterprises," Eurostat Statistics Explained (website), January 2021, https://ec.europa.eu/eurostat/statistics-explained/index.php/Cloud_computing_-_statistics_on_the_use_by_enterprises#Use_of_cloud_computing:_highlights.

128 Ariel E. Levite and Gaurav Kalwani, *Cloud Governance Challenges: A Survey of Policy and Regulatory Issues*, Carnegie Endowment for International Peace, November 9, 2020, https://carnegieendowment.org/2020/11/09/cloud-governance-challenges-survey-of-policy-and-regulatory-issues-pub-83124.

129 Herr, *Four Myths about the Cloud*.

130 Geranmayeh and Rapnouil, "Meeting the Challenge of Secondary Sanctions."

131 "US 'Cloud' Supremacy Has Europe Worried about Data," Euractiv (media network), August 4, 2020, https://www.euractiv.com/section/digital/news/us-cloud-supremacy-has-europe-worried-about-data/.

132 "Mass Surveillance: EU Citizens' Rights Still in Danger, Says Parliament," Press Release, European Parliament, October 29, 2015, https://www.europarl.europa.eu/news/en/press-room/20151022IPR98818/mass-surveillance-eu-citizens-rights-still-in-danger-says-parliament.

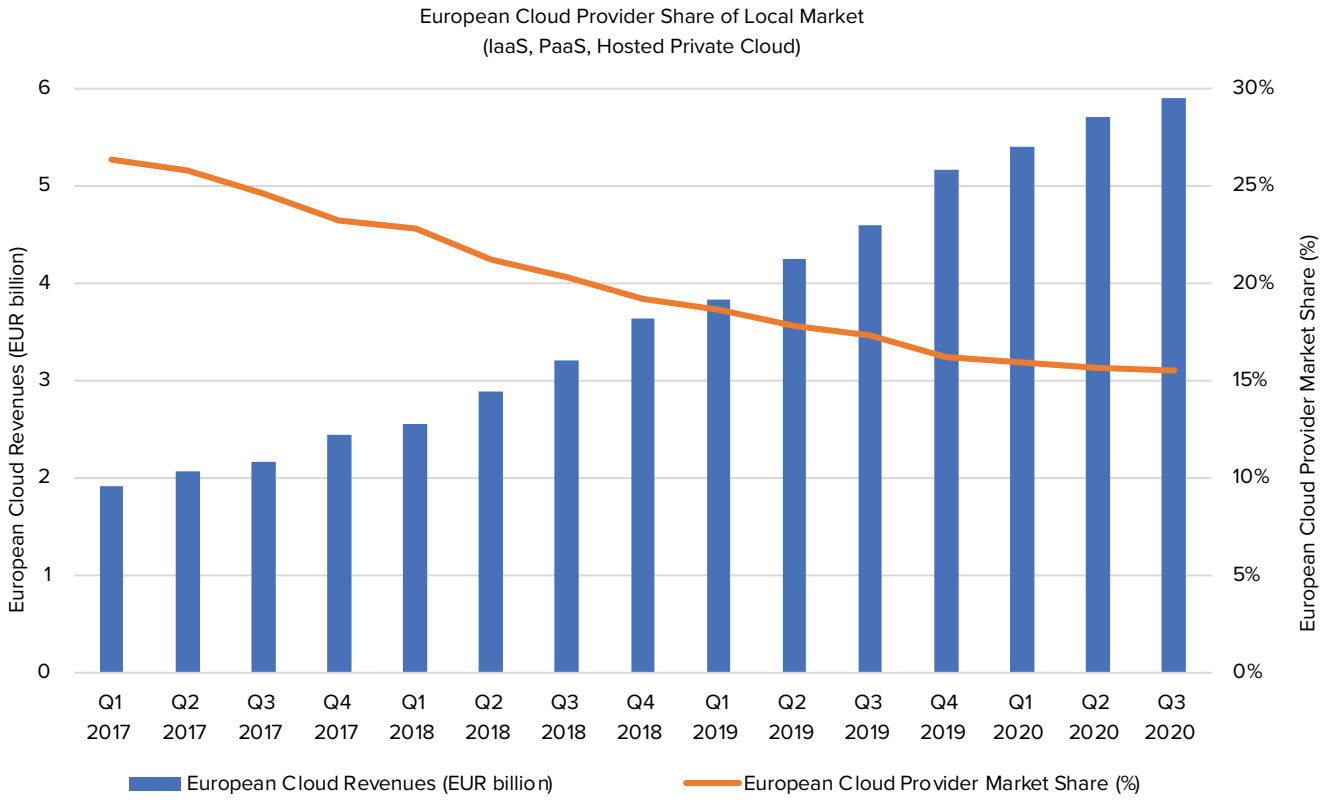
133 Laurens Cerulus, "Europe's Litmus Test over Cloud Computing Push," *Politico,* July 28, 2020, https://www.politico.eu/article/shades-of-sovereignty-dent-european-cloud-dreams/.

134 Janosch Delcker, "German Watchdog Says Amazon Cloud Vulnerable to US Snooping," *Politico,* April 4, 2019, https://www.politico.eu/article/german-privacy-watchdog-says-amazon-cloud-vulnerable-to-us-snooping/.

135 Foo Yun Chee, "EU Bodies' Use of Amazon, Microsoft Cloud Services Faces Privacy Probes," Reuters," May 27, 2021, https://www.reuters.com/article/ctech-us-eu-dataprotection-amazon-com-mi-idCAKCN2D80TQ-OCATC.

136 European Commission, *European Strategy for Data*.

Figure 3: Share of EU-based Cloud Providers on the European Market Declines[1]



Source: Synergy Research Group

1 "European Cloud Providers Struggle to Reverse Market Share Losses," Synergy.

However, the pursuit of strategic autonomy in cloud infrastructure and related data localization requirements have important limitations. Some EU member states such as France and Germany have already introduced data localization requirements that restrict the storage or transmission of certain data outside their borders.[137] However, as Atlantic Council research by Trey Herr indicates, data localization policies, especially when they do not adequately capture the different types of data managed by cloud operators, can undermine the public cloud's economic and technical underpinnings.[138] The public cloud—where a service provider makes its resources available to the public, enabling organizations and individuals to leverage large volumes of data storage and cutting-edge computing services at lower costs—is made possible by large numbers of global users accessing international networks of storage and computing resources.[139] Therefore, when data cannot cross borders, much of the flexibility at the core of the public cloud model can be lost.[140] Data localization requirements can force cloud service providers to build infrastructure to serve specific markets, potentially leading to costly changes in core infrastructure design.[141] For example, the push for data localization has led to a different model of cloud services provided by Microsoft Azure in Germany.[142] In 2015, Microsoft announced that it would open a cloud offering that would be run through a data trustee, which would control the data stored in Germany and Microsoft's access to the data.[143] However,

137 Herr, *Four Myths about the Cloud*.
138 Herr, *Four Myths about the Cloud*.
139 Herr, *Four Myths about the Cloud*.
140 Herr, *Four Myths about the Cloud*.
141 Herr, *Four Myths about the Cloud*.
142 Tim Maurer and Garrett Hinck, *Cloud Security: A Primer for Policymakers*, Carnegie Endowment for International Peace, August 31, 2020, https://carnegieendowment.org/2020/08/31/cloud-security-primer-for-policymakers-pub-82597.
143 Maurer and Hinck, *Cloud Security*.

this offering wasn't sustainable, according to reports, due to higher costs and separation from international business, which highlighted the challenges posed by data localization rules.[144] See chapter 2.2 for more analysis on data transfers and the impact of data localization rules.

National Considerations

The EU member states want to address dependency on non-European data infrastructure but are divided on how far to go in pursuit of strategic autonomy.[145] Some countries see both a risk to autonomy and privacy as well as an opportunity to advance European industry and help foster national "champions."[146] For example, the German and French led Gaia-X project aims to provide a platform to combine cloud-hosting services from several companies. It plans to achieve this by "allowing businesses to move their data freely with all information protected under Europe's tough data processing rules,"[147] and by doing so, aims to strengthen European digital sovereignty by decreasing dependencies on external digital infrastructure suppliers.[148] Yet some experts argue that it is not clear if the EU has sufficient domestic suppliers of cloud services to substantively replace foreign suppliers to the EU market.[149]

However, some EU member states are wary of potential protectionist measures that threaten Europe's openness to cloud services and favor larger member states.[150] In March 2020, Finland led a coalition of fifteen EU member states that published a joined statement asserting that, "to avoid distortions in the single market and to ensure a global level playing field, operators from third countries must compete under similar conditions and rules as their European counterparts."[151] This is potentially also because some member states are benefitting from US investment. For instance, in 2020 Poland saw major cloud infrastructure investments of up to $3 billion (combined) from Microsoft and Google.[152] Additionally, according to *Politico,* some European states are concerned that a push for greater European digital autonomy is going to provide the Franco-German industry a new edge at the expense of smaller economies.[153] In 2020, Finnish Minister for European Affairs Tytti Tuppurainen said: "There is a very strong tendency toward ever-growing protectionism."[154] The minister said she fears "smaller economies might be in danger of being trampled on."[155]

Business Considerations

Businesses need affordable data infrastructure that meets their need for fast analysis and interpretation of data. Cloud services can drive machine learning implementation because they provide organizations with access to high-performance infrastructure that they could not afford on their own.[156] For example, the rise of autonomous vehicles is fueled by the use of cloud service providers as they require large data storage, computing power, and real-time accessibility, with data traffic from autonomous vehicles predicted to increase a thousandfold between 2018 and 2025.[157] Edge computing infrastructure and standards also will be increasingly important for the deployment of IoT systems, for example in smart agriculture. When smart devices can share and process data more directly through edge computing—like when a sensor monitoring crops analyzes air moisture and directly triggers a watering system, rather than doing so via a server located hundreds of miles away—it can reduce the volume of data traffic needed to be transferred long distances, lowering costs, and enabling

144 Maurer and Hinck, *Cloud Security*.

145 Paola Tamma, "Europe Wants 'Strategic Autonomy'—It Just Has to Decide What That Means," *Politico*, October 15, 2020, https://www.politico.eu/article/europe-trade-wants-strategic-autonomy-decide-what-means/.

146 Tamma, "Europe Wants 'Strategic Autonomy' "; and Simon Van Dorpe, "Pressure Grows for Antitrust Action against German Software Giant," *Politico*, October 28, 2020, https://www.politico.eu/article/pressure-grows-for-antitrust-action-against-german-software-giant/.

147 Tamma, "Europe Wants 'Strategic Autonomy.' "

148 Delcker and Heikkila, "Germany, France Launch Gaia-X Platform."

149 Herr, *Four Myths about the Cloud*.

150 Tamma, "Europe Wants 'Strategic Autonomy.' "

151 "Strengthening the Economic Base of the EU," a Joint Statement of Fifteen Nations ahead of the March 2020 European Council, Permanent Representation of Finland to the EU, Finland Abroad (website), March 3, 2020, https://finlandabroad.fi/web/eu/current-affairs/-/asset_publisher/cGFGQPXL1aKg/content/joint-statement/384951.

152 Reuters staff, "Microsoft to Invest $1 Billion in Polish Cloud Project," Reuters, May 5, 2020, https://www.reuters.com/article/us-microsoft-poland-idUSKBN22H0WP; and Reuters staff, "Google to Invest up to $2 Billion in Polish Data Centre, Paper Says," Reuters, June 24, 2020, https://www.reuters.com/article/us-google-poland-idUSKBN23V0PA.

153 Tamma, "Europe Wants 'Strategic Autonomy.' "

154 Tamma, "Europe Wants 'Strategic Autonomy.' "

155 Pekka Vanttinen, "Helsinki: No to 'Fortress Europe' and Strategic Capitalism," Euractiv, October 8, 2020, https://www.euractiv.com/section/politics/short_news/helsinki-no-to-fortress-europe-and-strategic-capitalism/.

156 Anne Bonner, "Why Is Everybody Talking About the Cloud?," Medium (publishing platform), September 3, 2019, https://towardsdatascience.com/why-is-everybody-talking-about-the-cloud-b3fcfa1dda76.

157 Rich Millar, "For Autonomous Vehicles, The Road Ahead Is Paved with Data," *Data Center Frontier*, October 14, 2019, https://datacenterfrontier.com/for-autonomous-vehicles-the-road-ahead-is-paved-with-data/; and Rich Millar, "Rolling Zettabytes: Quantifying the Data Impact of Connected Cars," *Data Center Frontier*, January 21, 2020, https://datacenterfrontier.com/rolling-zettabytes-quantifying-the-data-impact-of-connected-cars/.

better service in areas with lower connectivity.[158] At the same time, edge computing presents security risks that require additional safeguards as this decentralized model expands the potential attack surface, making data harder to protect.[159]

The customers of cloud providers reap the benefits of economies of scale. The ability to access global networks of data storage and processing through public cloud offerings enables organizations to leverage large volumes of data storage and cutting-edge computing services at lower costs compared to in-house data centers.[160] However, this also leads to strong dependency on cloud services, which can create vulnerabilities for businesses.[161] For example, difficulties and costs of switching can exacerbate vendor lock-in, which could potentially lead to businesses having fewer and lower-quality service choices.[162]

Recommendations for the European Commission

A. Enhance the implementation of the European Alliance for Industrial Data, Edge and Cloud by seeking additional investments from the participating cloud computing providers and industrial cloud users.

While the project is still at its early stages, it would diversify the supply of cloud services available to European organizations, mitigating risks stemming from over-reliance on a small number of non-European providers. While the project received political support from the member states, the commission should seek additional incentives to enhance European industry players' involvement and financial contributions in the Federated Cloud project.

B. Enhance coordination between the European Alliance for Industrial Data, Edge and Cloud with parallel initiatives, such as Gaia-X, and ensure that the European Cloud Rulebook and other proposals do not impose significant regulatory burdens on cloud service providers that could impact the competitiveness of their offerings.

As argued by Andrea Renda and others from the Centre for European Policy Studies (CEPS), the commission should aim to maximize coordination with existing parallel projects, such as Gaia-X to minimize overlaps in rules and norms that would prevent sufficient interoperability, openness, and transparency of the EU data economy.[163] While there are not many details yet available regarding the Cloud Rulebook and marketplace proposals mentioned in the EU Data Strategy, the European Commission needs to ensure that the potential regulatory burden will not severely impact the ability of European providers to provide competitive price and service offerings.

2.2 Simplify data transfers

Training algorithms to automate tasks, make predictions, and develop new products and services often requires businesses to share data with machine learning specialists or to combine data sets across multiple locations.[164] Yet only 6 percent of European businesses are using data in this way.[165] With cross-border data flows now contributing more to global gross domestic product (GDP) than the flow of goods,[166] **the EU risks being excluded from global value chains due to its relatively slow uptake of data transfers.**

To address this, the European Commission should develop new mechanisms to enable EU-based businesses to participate in global data value chains and address legal ambiguity in EU legislation that increase risks and costs for businesses sharing data. One important issue related to data transfers in the EU is data protection regulations to safeguard the fundamental rights of European citizens. This chapter builds on chapter 2.1 and focuses on identifying barriers to data transfers and ways to address them. It is not intended to downplay the importance of maintaining

158 Renda et al., *The Digital Transition.*

159 Charles Owen-Jackson, "What Does the Rise of Edge Computing Mean for Cybersecurity?," Kaspersky (blog), 2019, https://www.kaspersky.com/blog/secure-futures-magazine/edge-computing-cybersecurity/31935/.

160 Ron Davies, *Cloud Computing: An Overview of Economic and Policy Issues: In Depth Analysis,* European Parliamentary Research Service, European Union, 2016, https://data.europa.eu/doi/10.2861/705354.

161 Levite and Kalwani, "Cloud Governance Challenges"; and "Cloud Computing: A Different Way of Using IT," European Union, 2019, https://digitalforeurope.eu/wp-content/uploads/2020/03/CloudComputingEC.pdf.

162 Levite and Kalwani, "Cloud Governance Challenges."

163 Renda et al., *The Digital Transition*.

164 Victoria Rees, "Ten Big Pharma Companies Collaborate on Data Sharing AI," *European Pharmaceutical Review* (June 7, 2019), https://www.europeanpharmaceuticalreview.com/news/89540/ten-big-pharma-companies-collaborate-on-data-sharing-ai/.

165 Martina Barbero et al., *Study on Emerging Issues of Data Ownership, Interoperability, (Re-)Usability and Access to Data, and Liability*, Prepared for the European Commission Directorate-General of Communications Networks, Content and Technology by Deloitte, April 25, 2018, 31, doi: 10.2759/781960.

166 James Manyika et al., *Digital Globalization: The New Era of Global Flows*, McKinsey & Company's McKinsey Global Institute, March 2016, https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20globalization%20The%20new%20era%20of%20global%20flows/MGI-Digital-globalization-Full-report.pdf.

protections, particularly for personal data, and should be read in close conjunction with chapter 3.1 on enhancing privacy and trust.

EU Policy Context

Within the EU, aspects of European legal frameworks lack clarity on data transfer issues.[167] While the European Commission has funded the Support Center for Data Sharing to help organizations overcome legal barriers,[168] there is a need for greater clarity in relevant legislation.[169] The proposed Data Governance Act introduces a legal framework for "data sharing services" aimed at increasing trust and reducing costs to data sharing,[170] however the current proposal does not address the legal uncertainties that give rise to those costs.[171]

The Schrems II judgment in the EU Court of Justice invalidated widely used practices for sharing data outside the EU. The GDPR required that personal data transferred outside the EU must not undermine the protection set by the regulation.[172] This condition can be met by a European Commission "adequacy decision" determining that a country provides an adequate level of data protection.[173] However, the July 2020 Schrems II judgment invalided such an "adequacy decision" negotiated with the United States under the EU-US Privacy Shield—an agreement used by more than 5,300 companies.[174] The judgment also brought into question the transfer of data under standard contractual clauses, which are commonly used data-protection obligations.[175] As a result, such clauses are subject to onerous new safeguards that can be impractical for data exporters.[176] This has heavily impacted businesses that transfer data, with about 85 percent of EU-based businesses that share data previously using such clauses to transfer data outside the EU.[177] A study commissioned by the European Parliament indicates that renewing the adequacy decision with the United States will likely require fundamental changes to the US data privacy and surveillance regimes.[178]

Geopolitical Considerations

Cross-border data transfers can expose sensitive data to weaker data protection regimes and make it easier for foreign governments to surveil and influence European citizens. Further to Chapter 2.1, which includes a focus on the risks of surveillance through foreign-owned data infrastructure, data transferred to a non-EU jurisdiction can be subject to foreign surveillance and manipulation. (These risks are in addition to those created by relying on foreign-owned cloud services, which become part of critical infrastructure, and are vulnerable to disruption.) For example, in 2008 the US Congress amended the Foreign Intelligence Surveillance Act to allow US intelligence agencies to conduct surveillance of non-US persons located abroad without individual warrants under Section 702.[179] Once general targeting is authorized by the Foreign Intelligence Surveillance Court, the US government can require electronic communication service providers, like email, telephone, and remote computer service providers, to provide information and assistance, including information in US-based facilities such

167 Amended Proposal of the Commission of the European Communities for the European Parliament and Council Regulation on the Law Applicable to Non-Contractual Obligations ("Rome II") (February 21, 2006), Regarding Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products, 85/374/EEC (July 25, 1985), Pub. L. No. 31985L0374, *Official Journal of the European Union*, L210 (1985), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52006PC0083&qid=1633805012772.

168 Support Centre for Data Sharing website, provided and managed by the European Commission, Directorate-General for Communications Networks Content and Technology, accessed April 18, 2021, https://eudatasharing.eu/homepage.

169 European Commission, "Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics," Submitted to the European Parliament, the Council, and the European Economic and Social Committee, February 19, 2020, accessed April 17, 2021, https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf.

170 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act).

171 "Towards a New Data-Sharing Ecosystem across the EEE: The Data Governance Act," Osborne Clark (law firm website), January 25, 2021, https://www.osborneclarke.com/insights/towards-new-data-sharing-ecosystem-across-eee-data-governance-act/.

172 "General Data Protection Regulation (GDPR)–Official Legal Text," (EU) 2016/679, chap. V, Intersoft Consulting (website), accessed April 18, 2021, https://gdpr-info.eu/.

173 W. Gregory Voss, "Cross-Border Data Flows, the GDPR, and Data Governance," *Washington International Law Journal* 29, no. 3 (June 2020): 485–532, https://hal.archives-ouvertes.fr/hal-02872471.

174 Kenneth Propp and Peter Swire, "Geopolitical Implications of the European Court's Schrems II Decision," *Lawfare* (blog), Lawfare Institute in coordination with Brookings Institution, July 17, 2020, https://www.lawfareblog.com/geopolitical-implications-european-courts-schrems-ii-decision.

175 Anupam Chander, "Is Data Localization a Solution for Schrems II?" *Journal of International Economic Law*, forthcoming, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3662626.

176 Ira Rubinstein and Peter Margulies, "Risk and Rights in Transatlantic Data Transfers: EU Privacy Law, US Surveillance, and the Search for Common Ground," Social Science Research Network (SSRN) Scholarly Paper, February 16, 2021, 21–32, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3786415.

177 Digital Europe, BusinessEurope, European Round Table for Industry (ERT), and the European Automobile Manufacturers Association, "Schrems II Impact Survey Report," November 26, 2020, https://www.digitaleurope.org/resources/schrems-ii-impact-survey-report/.

178 Ian Brown and Douwe Korff, "Exchanges of Personal Data After the Schrems II Judgment," n.d., 8–13.

179 Rachel Miller, "FISA Section 702: Does Querying Incidentally Collected Information Constitute a Search Under the Fourth Amendment?," *Notre Dame Law Review Reflection* 95, no. 3 (March 1, 2020), 139, https://scholarship.law.nd.edu/ndlr_online/vol95/iss3/2.

as data centers.[180] China has put in place similar foreign intelligence surveillance laws, such as the 2017 National Intelligence Law which obliges individuals, organizations, and institutions to assist national security agencies in "intelligence" work.[181] Beyond surveillance, inadequate data protections deployed by businesses outside the EU have allowed bad actors, like Russia, to unleash disinformation campaigns and influence elections in Europe.[182] These instances are among some of the drivers of EU policies to implement legal restrictions on cross-border data transfers. They also serve to support calls for protectionist data localization policies intended to shelter domestic industry from international competition.[183]

However, flows of data outside the EU are critical for global collaboration, economic competitiveness, and innovation. The Digital Europe survey found that of the EU-based businesses transferring data to another nation, 94 percent of them send data to the United States, 56 percent to the United Kingdom, and 59 percent to an Asian nation.[184] These data transfers underpin essential business functions and are a driving force behind global-value chains—where business operations are divided across countries for efficiency, lower costs, and faster production.[185] As cross-border data flows increasingly become an integral part of the global economy,[186] countries that participate in a greater share of these value chains will see more economic benefits from the digital economy.[187] While the commission still has in place adequacy agreements with fourteen countries,[188] surveys (conducted before the June 2021 agreement with the United Kingdom) indicate these agreements are only used by 5 percent of EU businesses transferring data.[189] To secure its place in future global value chains, the EU needs mechanisms in place that seamlessly facilitate data transfers with global innovation hubs, particularly the United States.

National Considerations

EU member states that are more concerned about security and data privacy are in favor of restricting cross-border data flows. As detailed in chapter 2.1, there are calls in some EU member states for data localization rules that restrict the storage or transmission of certain data outside their borders**.** The proposal for France's Digital Republic Bill (2016) included a provision to completely restrict data transfers outside the EU.[190] Although the provision was later removed, it is representative of the desire among many French and European leaders to defend Europe's "digital sovereignty." This includes both the protection of data from foreign governments and businesses, as well as maintaining the ability for European law enforcement to access European data when needed.[191] For example, European law enforcement agencies have been given some leeway under the 2020 *Quadrature du Net and Others decision* of the Court of Justice of the European Union to intercept and retain bulk data indiscriminately when faced with "serious threats to national security," such as terrorist activities.[192]

But smaller, more open and trade-dependent member states want fewer barriers to cross-border data flows. What had been dubbed the Digital Nine (D9) group of EU member states wants the EU to be more active in pursuing more open cross-border data flows.[193] In a 2017 letter addressed to the Vice President of the European Commission, the group urged the EU to include ambitious rules on open data flows in trade agreements, arguing

180 Chris D Linebaugh and Edward C Liu, "EU Data Transfer Requirements and US Intelligence Laws: Understanding Schrems II and Its Impact on the EU-US Privacy Shield," n.d., 9.

181 Murray Scot Tanner, "Beijing's New National Intelligence Law: From Defense to Offense," *Lawfare* (blog), Lawfare Institute with the Brookings Institution, July 20, 2017, https://www.lawfareblog.com/beijings-new-national-intelligence-law-defense-offense.

182 Bignami, "Schrems II."

183 Alan Beattie, "Data Protectionism: The Growing Menace to Global Business," *Financial Times*, May 13, 2018, https://www.ft.com/content/6f0f41e4-47de-11e8-8ee8-cae73aab7ccb.

184 Digital Europe, BusinessEurope, ERT, and the European Automobile Manufacturers Association, "Schrems II Impact Survey Report."

185 Magnus Rentzhog et al., *No Transfer, No Trade*.

186 Manyika et al., *Digital Globalization*.

187 Silja Baller, Soumitra Dutta, and Bruno Lanvin, eds., *The Global Information Technology Report 2016: Innovating in the Digital Economy*, World Economic Forum and INSEAD, 2016, 40, https://www.insead.edu/sites/default/files/assets/dept/globalindices/docs/GITR-2016-report.pdf.

188 "Adequacy Decisions: How the EU Determines If a Non-EU Country Has an Adequate Level of Data Protection," European Commission (website), accessed April 17, 2021, https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

189 Digital Europe, BusinessEurope, ERT, and the European Automobile Manufacturers Association, "Schrems II Impact Survey Report."

190 Olivier Proust, "France Adopts Digital Republic Law," Fieldfisher (website of European law firm), dated October 4, 2016, accessed 2021, https://www.fieldfisher.com/en/services/privacy-security-and-information/privacy-security-and-information-law-blog/france-adopts-digital-republic-law.

191 Winston Maxwell and Mathilde Gerot, "Ignoring GDPR, French Senate Votes for a Data Localization Amendment," *HL Chronicle of Data Protection* (blog now called Hogan Lovells Engage), Hogan Lovells (law firm) June 1, 2016, https://www.lexology.com/library/detail.aspx?g=3767ab61-1ab3-403f-882d-d3f7fc794dfb.

192 La Quadrature Du Net and Others v Premier Ministre and Others, ECJ Grand Chamber, ECLI:EU:C:2020:791, (2020), paragraphs 135–138, Curia (case-law database), accessed August 12, 2021, https://curia.europa.eu/juris/document/document.jsf?docid=232084&doclang=en.

193 When formed in 2016, the D9 included Sweden, Denmark, Finland, Estonia, Belgium, the Netherlands, Luxembourg, Ireland, and the United Kingdom, which has since left the EU. See "Ann Linde to Host Digital 9 Ministerial Meeting," Government Offices of Sweden (website), October 17, 2017, https://www.government.se/press-releases/2017/10/ann-linde-to-host-digital-9-ministerial-meeting/.

that the absence of those rules "limits [their] countries, citizens, and entrepreneurs from participating in global value chains."[194] In March 2021, a similar group of eight smaller EU member states co-signed a letter underlining the need to "eliminate barriers to cross-border online services, and to ensure free data flows."[195] However, the European Data Protection Supervisor has emphasized in relation to negotiations with the United Kingdom (UK), that data protection is nonnegotiable in international trade agreements, which must preserve individuals' fundamental rights to data protection and privacy.[196] While the commission did reach an adequacy decision with the UK, for the first time it included a sunset clause which means the decision will automatically expire in four years.[197]

Business Considerations

The lack of legal clarity on data ownership, liability, and reuse creates costs and risks to businesses sharing data. Companies rely on contracts when sharing data. For example, about 60 percent of the EU-based businesses that share data do so with subcontracted data analytics services.[198] However, legal uncertainties mean data-sharing contracts often need to be tailored to address specific issues and examined on a case-by-case basis,[199] requiring complex and expensive negotiations. Liability issues are one example, given the risks when sharing data with a third-party whose misuse of the data could result in damages.[200] Similarly, selling a product or service based on someone else's data could lead to damages and harm if the data is incorrect or significantly biased (see chapter 3.2 for more on bias and data-driven discrimination).

Transferring data outside the EU is even more complex, costly, and risky. While the Schrems II judgment applies to personal data, most data sets contain both personal and non-personal data,[201] and it is often impractical, if not impossible, to split mixed data sets.[202] This means that the GDPR and the Schrems II judgment apply to a large proportion of data sets in the EU. Following the judgment, Digital Europe conducted a business survey that showed that 91 percent of EU-based businesses that share data do so with entities outside the EU and 90 percent reported facing moderate-to-high costs in reassessing their data-sharing contracts to comply with the judgment.[203] These costs mean smaller businesses are more likely to avoid cross-border data transfers at all.[204]

Recommendations for the European Commission

A. Give greater legal clarity on data ownership and on who is responsible for damages that occur due to externally provided data.[205] This includes in relation to:

i. The General Product Safety Directive not containing any provisions on the liability of damages caused by data where products are based on third-party data providers, which creates a barrier for the development of such business models.[206]

ii. The Product Liability Directive's ambiguous scope in relation to digital products and defects resulting from errors in externally provided data.[207]

iii. The need for legal instruments that define data ownership and data access/use rights for nonpersonal data (see recommendation B in chapter 2.3).

Clarifying liability and ownership issues in EU-wide legal frameworks reduces the need for these issues

194 Joint Letter of Ten EU Countries to Frans Timmermans on Data Flows and Data Localization, *Politico* (scan), May 15, 2017, https://www.politico.eu/wp-content/uploads/2017/05/POLITICO_joint-letter-eu-countries-data-flows-in-trade_May-15-2017.pdf?utm_source=POLITICO.EU&utm_campaign=38a66c58b5-EMAIL_CAMPAIGN_2017_05_12&utm_medium=email&utm_term=0_10959edeb5-38a66c58b5-190045157.

195 "The President: Progress towards Creating a Single Digital Market Is Key for EU's Economic Recovery," President of the Republic of Lithuania (website), accessed August 13, 2021, https://www.lrp.lt/en/media-center/news/the-president-progress-towards-creating-a-single-digital-market-is-key-for-eus-economic-recovery/35554.

196 "Data Protection Is Non-Negotiable in International Trade Agreements," European Data Protection Supervisor (website), Press Release of February 22, 2021, accessed August 6, 2021, https://edps.europa.eu/press-publications/press-news/press-releases/2021/data-protection-non-negotiable-international_en.

197 "Data Protection: Commission Adopts Adequacy Decisions for the UK," European Commission (website), Press Release of June 28, 2021, accessed August 6, 2021, https://ec.europa.eu/commission/presscorner/detail/en/ip_21_3183.

198 Barbero et al., *Study on Emerging Issues of Data Ownership*.

199 Barbero et al., *Study on Emerging Issues of Data Ownership*, 77 and 82.

200 Barbero et al., *Study on Emerging Issues of Data Ownership*, 81–82.

201 "Commission Publishes Guidance on Free Flow of Non-Personal Data: Questions and Answers," European Commission (website), May 29, 2019, accessed 2021, https://ec.europa.eu/commission/presscorner/detail/en/MEMO_19_2750.

202 Stéphanie De Smedt, "Free Flow of Non-Personal Data and GDPR," Loyens & Loeff (law firm website), June 19, 2019 (post), accessed 2021, https://www.loyensloeff.com/en/en/news/news-articles/free-flow-of-non-personal-data-and-gdpr-n14929/.

203 Digital Europe, BusinessEurope, ERT, and the European Automobile Manufacturers Association, "Schrems II Impact Survey Report."

204 Chander, "Is Data Localization a Solution for Schrems II?"

205 This point was raised in the Deloitte review of data and liability in the EU. See Barbero et al., *Study on Emerging Issues of Data Ownership*.

206 Barbero et al., *Study on Emerging Issues of Data Ownership*, 84.

207 Barbero et al., *Study on Emerging Issues of Data Ownership*, 119.

to be addressed in contractual arrangements when businesses share data and lowers the uncertainty and risk in legal proceedings. This work would particularly support innovative business models among small and medium-sized enterprises (SMEs) that cannot afford the potential legal costs of these business models under the current framework.[208] This also would help support a high-growth sector in the EU that currently feels neglected by policy makers in Brussels.[209]

B. Prioritize negotiations on the recognition of data protection adequacy with global innovation hubs that share European values on data protection to increase the competitiveness of EU businesses in global value chains. This should involve exploring solutions with the United States to address the risks highlighted in Schrems II, while enabling cross-border economic activity.[210]

Data transfers with the United States and other innovation hubs are embedded in the value chains of many EU businesses, and value chains globally. Establishing mechanism to seamlessly transfer data with adequate protections would strengthen the position of European companies in the global economy. Adequacy determinations are also often preceded by substantial discussions with the third country, which in some cases have resulted in changes to data protection practices or even legislation in third countries to reflect European standards more closely.[211] This adoption of European standards increases the odds that emerging global digital norms more closely align with European systems and values.

Finding ways to reconcile EU data protection and privacy requirements with US surveillance laws and its lack of a national data privacy law will continue to pose significant challenges to an adequacy decision and likely require significant changes to US data privacy and surveillance regimes.[212] That said, there are ideas being put forward that could potentially offer solutions to the impasse, such as the hybrid model offered by Rubinstein and Margulies, which is based on US export control law.[213]

2.3 Foster inclusive data ecosystems

Like public-sector data (see chapter 1.1), making private data sets more widely available allows other organizations and individuals to leverage them for machine learning. However, a 2018 study by Deloitte found that European companies often cannot access the data they need from other companies, and when they do, face strict contractual limitations on using it.[214] It also found that smaller companies and those in a weaker position in the value chain had unequal bargaining power when seeking to access data.[215] This lack of data access can put businesses at a disadvantage in the market, stop them from developing new products and services, and slow productivity and efficiency growth.[216] While the problem varies across sectors, the lack of competition linked to the unwillingness of companies to share data—alongside other impediments to data access—has negative implications for the EU's digital competitiveness and innovation, freedom of choice for consumers, and digital inclusion.[217] **The EU needs to encourage data sharing and remove anticompetitive data barriers while maintaining incentives for businesses to invest in data.**

To achieve this, the European Commission should adopt sector-specific data portability rights and introduce rules on when businesses should allow access to proprietary data sets to promote competition in European data ecosystems. The commission should also establish sector-specific consultative mechanisms to review these as technologies and market structures change over time.

This chapter should be read in conjunction with chapter 3.1 on privacy and trust for its greater focus on protecting the rights of individuals, which is fundamental when sharing data between organizations.

EU Policy Context

The commission has been developing ways to incentivize businesses to make data more widely available. This includes data portability rights for customers—the ability to extract your data from one service and use it

208 Barbero et al., *Study on Emerging Issues of Data Ownership*, 84.

209 Pieter Haeck, "European Startups Are Booming. So Why Is Brussels Still Obsessed with Big Tech?," *Politico*, August 3, 2021, https://www.politico.eu/article/pr-problem-high-flying-startups-not-brussels-priority/.

210 Rubinstein and Margulies, "Risk and Rights in Transatlantic Data Transfers."

211 Ignacio Garcia Bercero and Kalypso Nicolaidis, *The Power Surplus: Brussels Calling, Legal Empathy and the Trade-Regulation Nexus*, Technical Report, Centre for European Policy Studies (CEPS) Policy Insights, PI 2021/05, School of Transnational Governance (STG) Report, 2021, 9, accessible via European University Institute Research Repository, https://hdl.handle.net/1814/70675.

212 Brown and Korff, "Exchanges of Personal Data After the Schrems II Judgment," 8–13.

213 Rubinstein and Margulies, "Risk and Rights in Transatlantic Data Transfers."

214 Barbero et al., *Study on Emerging Issues of Data Ownership*, 15.

215 Barbero et al., *Study on Emerging Issues of Data Ownership*, 15.

216 Barbero et al., *Study on Emerging Issues of Data Ownership,* 387–88.

217 Barbero et al., *Study on Emerging Issues of Data Ownership*, 16, 46.

for another—such as the 2016 General Data Protection Regulation, which gives individuals the right to port personal data.[218] The commission is also implementing sector-specific rules on access to proprietary data sets—the right for a business to access specific data held by others—through the proposed Digital Markets Act, which puts obligations on large platforms to give businesses real-time access to certain data related to the businesses' products and services.[219] The commission has further foreshadowed sectorial legislation to create "common data spaces" that encourage, and possibly even force, entities to make their data more widely available.[220] It is also preparing a Data Act, which aims to create a "fair data economy," including by clarifying rights related to nonpersonal data.[221]

Geopolitical Considerations

Major powers want access to data that helps their industries compete internationally. European leaders are concerned that the strategic acquisition of companies at the center of data ecosystems could be leveraged for competitive advantage over EU-based industries.[222] For example, China's "Made in China 2025" policy to develop high-tech manufacturing encouraged Chinese firms to strategically acquire foreign companies with advanced technologies and data.[223] In a conversation with digital policy expert Jörn Fleck from the Atlantic Council, he observed that European leaders may fear a repeat of the mid-2000s when the EU "missed the boat" on digital platforms, resulting in US firms dominating this space; however, he noted that this time it would be worse. With traditional industries like manufacturing turning to machine learning to boost productivity, a lack of control over data in the value chain could weaken Europe's industrial base, he added.

Policies designed to redistribute data in favor of European firms can pose risks to the EU. Comments by European Internal Markets Commissioner Thierry Breton—"European data will be used for European companies in priority, for us to create value in Europe"—suggest plans to create data ecosystems that treat European companies more favorably.[224] However, in a conversation with Andrea Renda, a digital economy expert at CEPS, he warned that while Europe's desire to develop its own solutions is legitimate and desirable, policies designed to insulate European tech companies from outside competition would weaken the EU's competitiveness in the digital economy, and consequently, the EU's share of the global economy as it becomes increasingly digitized. Furthermore, such policies could invite retaliation from trading partners. Although unrelated to data, US threats to impose tariffs if EU member states enact a digital services tax on platforms such as Facebook, Google, and, Amazon demonstrate its readiness to impose costs on EU businesses when the interests of its big tech companies are harmed.[225]

National Considerations

EU rules encouraging or obliging businesses to share data need to consider the varying positions of member states. France has called on regulation that encourages platforms to share data to be "agile and flexible" to adapt to changing business models, while Germany seeks stronger rules to prevent anticompetitive behavior before it occurs.[226] This includes Germany's recent amendments to its national competition law to prevent dominant companies from using data to make market entry more difficult for other companies and to stop them from hampering data portability and interoperability.[227] Poland broadly agrees with the German approach, though it emphasizes the need for only a narrow catalog of black-listed practices, covering only those most harmful to competition that should be determined by member states.[228] Finland stresses that any new rules obliging companies to share data should

218 "General Data Protection Regulation (GDPR)," (EU) 2016/679 (2016), https://gdpr-info.eu/.

219 European Commission, DG CNECT, Proposal for a Regulation (Digital Markets Act), 2020/0374/COD, sec. 6(i).

220 European Commission, "Towards a Common European Data Space," Communication to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions, April 25, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232&from=EN.

221 "Data Act & Amended Rules on the Legal Protection of Databases," European Commission (website), accessed September 15, 2021, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-including-the-review-of-the-Directive-96-9-EC-on-the-legal-protection-of-databases-/public-consultation_en.

222 "Speech by Commissioner Thierry Breton at Hannover Messe," European Commission (website), July 15, 2020, https://ec.europa.eu/commission/presscorner/detail/es/speech_20_1362.

223 Rishav Gupta, "Made in China 2025, Acquisitions of Companies in EU by Chinese Companies," Nepal Economic Forum, November 21, 2018, https://nepaleconomicforum.org/uncategorised/made-in-china-2025-acquisitions-of-companies-in-eu-by-chinese-companies-and-its-implications/.

224 Burwell and Propp, *The European Union and the Search for Digital Sovereignty*.

225 David Lawder, "US Trade Chief Readies Tariffs against Six Countries over Digital Taxes," Reuters, March 27, 2021, https://www.reuters.com/article/us-usa-trade-digital/u-s-trade-chief-readies-tariffs-against-six-countries-over-digital-taxes-idUSKBN2BI31K.

226 Samuel Stolton, "Digital Brief, Powered by Google: DSA and DMA – Member States Respond," Euractiv, December 18, 2020, https://www.euractiv.com/section/digital/news/digital-brief-powered-by-google-dsa-and-dma-member-states-respond/.

227 "Amendment of the German Act against Restraints of Competition," Federal Cartel Office of Germany (website and press release of national competition regulator), January 19, 2021, https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2021/19_01_2021_GWB%20Novelle.html.

228 "Digital Services Regulations in the EU–Our Position," Republic of Poland, Digital Affairs–Chancellery of the Prime Minister (website), September 15, 2020, https://www.gov.pl/web/digitalization/digital-services-regulations-in-the-eu--our-position.

be "justified, proportionate, and clear," avoid "unnecessary administrative burden," and provide legal protection for platforms that control data.[229]

Business Considerations

Incumbent businesses seek to protect their data to give them a competitive advantage. Research shows that many incumbents in the market are not willing to share data due to a fear of new competition.[230] One reason for this is the ability to maintain a dominant position in the market through data network effects and machine learning.[231] For example, as technology analyst Benedict Evans explains,[232] autonomous vehicle (AV) manufacturers can use exclusive access to data to outperform competitors. This is because AVs capture images as they drive, comparing them with prebuilt 3D models of the vehicles' surroundings to understand points of difference using machine learning. As the vehicle drives, it is updating the model so that the next time it drives down the same road, there are fewer points of difference for it to process and hence a lower chance of it becoming confused and reacting dangerously. When this data is shared between vehicles, each one is able to drive more safely as the models are updated more frequently. By sharing this mapping data, AV manufacturers can help all AVs on the road to perform better and safer for the public. However, companies in a dominant position with more cars on the road are incentivized to withhold the data from competitors to distinguish themselves by offering a superior product.[233]

Data sharing obligations help create data ecosystems but can also have unintentional consequences. Data portability rights can create opportunities for upstream businesses and value for consumers including for vehicles repair services,[234] and pay-how-you-drive insurance.[235] Supplier access to in-vehicle data can also help downstream businesses such as in-built car component manufacturers improve their products to develop innovations like machine learning systems that forecast parts failures and recommend repairs.[236] However, competition experts warn that excessive obligations to share data might create disincentives for companies to invest in high-quality data sets.[237] For instance, there may be fewer opportunities to make a return on the investment,[238] or it could result in anticompetitive information exchange that puts them at a disadvantage.[239] The 2018 study by Deloitte found that these issues are unique to each sector due to the different business models, public interests, and economic sensitivities.[240] The case outlined above is one example where the public could potentially be better served by autonomous vehicle companies sharing mapping data to improve the safety of all vehicles on the road; however, in other cases, sharing data could disincentivize innovation or benefit certain consumers at the expense of others,[241] with little or no public benefit. The commission needs to carefully calibrate measures sector-by-sector to avoid outcomes that disproportionally impact certain groups, create barriers to competition, or hamper innovation.

Recommendations for the European Commission

In forthcoming sectorial legislation to create common data spaces in manufacturing, environment protection, mobility, health, financial, energy, agriculture, public administration, and skills,[242] the commission should:

A. Adopt sector-specific data portability rights and standards for customer data (personal and nonpersonal) that consider the distinct data-access problems at an industry level to balance the benefits and costs in different contexts.

229 "Finland Supports Common EU Rules on Digital Platforms," Press Release, Finnish Ministry of Economic Affairs and Employment (website), April 2, 2021, https://tem.fi/en/-/finland-supports-common-eu-rules-on-digital-platforms.

230 Barbero et al., *Study on Emerging Issues of Data Ownership*, 79.

231 Matt Turck, "The Power of Data Network Effects," Blog, January 4, 2016, https://mattturck.com/the-power-of-data-network-effects/.

232 Benedict Evans, "Winner-Takes-All Effects in Autonomous Cars," Independent Analysis, August 22, 2017, https://www.ben-evans.com/benedictevans/2017/8/20/winner-takes-all.

233 European Commission, "Digital Markets Act."

234 Bertin Martens and Frank Mueller-Langer, "Access to Digital Car Data and Competition in Aftermarket Maintenance Services," *Journal of Competition Law & Economics* 16, no. 1 (April 30, 2020): 116–41, https://doi.org/10.1093/joclec/nhaa005.

235 Dimitrios I. Tselentis, George Yannis, and Eleni I. Vlahogianni, "Innovative Insurance Schemes: Pay As/How You Drive," *Transportation Research Procedia*, Transport Research Arena TRA2016, 14 (January 1, 2016): 362–71, https://doi.org/10.1016/j.trpro.2016.05.088.

236 Dave Leggett, "How Machine Learning Can Forecast Parts Failures," November 28, 2018, Just Auto (industry portal), https://www.just-auto.com/interview/how-machine-learning-can-forecast-parts-failures_id185857.aspx.

237 Jordi Casanova, "Online Search Competition and the Risk of Unintended Consequences of Data Access," Competition Policy International (website), November 26, 2020, accessed 2021, https://www.competitionpolicyinternational.com/online-search-competition-and-the-risk-of-unintended-consequences-of-data-access/.

238 Jay Modrall, "Big Data, Big Target for EU Antitrust Enforcement?," Competition Policy International (website), February 24, 2020, accessed 2021, https://www.competitionpolicyinternational.com/big-data-big-target-for-eu-antitrust-enforcement/.

239 Modrall, "Big Data, Big Target."

240 Barbero et al., *Study on Emerging Issues of Data Ownership*, 69.

241 Barbero et al., *Study on Emerging Issues of Data Ownership*, 69.

242 European Commission, "A European Strategy for Data," COM(2020), 66.

The commission could draw lessons from the Australian Consumer Data Right, which is a series of sector-specific legislation giving consumers and small businesses control over their personal and nonpersonal consumer data in phases, starting with banking (2019), energy (2020), and telecommunications (forthcoming).[243]

B. Adopt sector-specific data-access obligations for platforms and dominant companies to share data with other businesses that consider the business models, public interest, and economic sensitivities of each sector.

This would limit companies from withholding access to data needed by other companies to innovate and compete, without unfairly discriminating against incumbent players.

C. Establish sector-specific consultative mechanisms involving industry, the proposed European Data Innovation Board, national standards organizations, and competition authorities to periodically review the suitability of sectoral data-access rules as technologies and market structures change.

For example, the Japanese Fair Trade Commission has created the Study Group on Competition Policy in Digital Markets, which includes experts from government, industry, and academia and meets monthly to investigate issues related to Japan's Antimonopoly Act.[244]

Implementing sector-specific portability and data-access rules would help balance European economic autonomy and openness by enabling a greater diversity of suppliers, particularly in downstream services, making Europe more resilient against geographic supply disruptions, and increasing opportunities for SMEs to participate in the value chain. Periodically reviewing these rules mitigates against "set and forget" policies that over time risk disincentivizing innovation and disadvantaging certain players owing to the changing nature of technology and digital markets.

243 "Consumer Data Right," Treasury Department of the Australian Government (website), accessed 2021, https://treasury.gov.au/consumer-data-right.

244 "The Study Group on Competition Policy in Digital Markets," Japan Fair Trade Commission (website), Press Release of July 22, 2020, https://www.jftc.go.jp/en/pressreleases/yearly-2020/July/200729.html.

Part Three: Protect Fundamental Rights

The European Commission anchors its approach to AI policies around the need to ensure the protection of fundamental rights while boosting research and industrial capacity.[245] Data rules for machine learning have the potential to directly infringe on citizens' private lives and the right to the protection of personal data (Articles 7 and 8 of the EU Charter of Fundamental Rights) and on nondiscrimination and equality between women and men (Articles 21 and 23).[246] Therefore, Chapters 3.1 and 3.2 focus on issues related to risks to privacy and discrimination enabled by data for machine learning.[247]

3.1 Enhance privacy and trust

To enable sustainable development and adoption of machine learning, the EU needs to reconcile privacy considerations with the benefits of access and reuse of data. The costs and legal difficulties of meeting strict privacy restrictions reduces the amount of data available to businesses and researchers for machine learning and can stymie European innovation.[248] At the same time, EU decision makers face important societal considerations as many Europeans lack confidence in organizations' and governments' ability and willingness to protect their data, despite the fact that the EU has one of the world's strictest data privacy regulations—the GDPR.[249] **Failure to protect citizens' fundamental rights, including privacy, while enabling wider data use could further undermine citizens' trust, European innovation and international competitiveness, and, consequently, the credibility of the EU's global values-based leadership.**

To address this, the European Commission should clarify guidance on applying the GDPR to machine learning techniques to diminish the costs of compliance and legal uncertainty while enhancing data-protection enforcement. The commission should also invest in emerging privacy-preserving machine learning techniques that can still produce accurate machine learning outputs while reducing the reliance of organizations on large, centralized sets of personal data to innovate. It should enhance research collaboration with like-minded countries in developing these techniques.[250]

EU Policy Context

The EU is struggling to balance implementing strict data protection with enabling data-driven innovation. The GDPR harmonized data-privacy laws across EU member states and overhauled how organizations process and handle data. It introduced limits to the full deployment of machine learning, for example by putting restrictions on certain automated decisions.[251] At the same time, the 2020 Data Governance Act (DGA) proposal aims to foster the availability of data for use and data-sharing mechanisms across the EU.[252] For example, the act seems to imply that public-sector bodies will be obliged to fulfill requests for access to personal data that has been anonymized.[253] However, this has been criticized by digital-rights organizations as anonymization techniques may not be sufficient to guarantee that data will not be traced back and reidentified.[254] These two policy examples show the difficulty of balancing the complex and sometimes competing considerations of privacy and innovation.

245 European Commission, "A Europe Fit for the Digital Age: Commission Proposes New Rules and Actions for Excellence and Trust in Artificial Intelligence," Press Release of April 21, 2021, accessed June 1, 2021, https://s3platform.jrc.ec.europa.eu/documents/portlet_file_entry/20125/Europe+fit+for+the+Digital+Age+Commission+proposes+new+rules+and+actions+for+excellence+and+trust+in+Artificial+Intelligence_21.04.21.pdf/bc6b6257-082a-a9fc-9c8c-f3e2e7c80ac4.

246 Nondiscrimination and equality pertaining to, for example, sex, race, color, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age, or sexual orientation.

247 Consolidated Version of the Treaty of the European Union, *Official Journal of the European Union*, 2016/C 202/01 (2016), https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12016ME/TXT&from=EN); and Charter of Fundamental Rights of the European Union, *Official Journal of the European Union*.

248 Erik Brattberg, Raluca Csernatoni, and Venesa Rugova, *Europe and AI: Leading, Lagging Behind, or Carving Its Own Way?*, Carnegie Endowment for International Peace, July 9, 2020, https://carnegieendowment.org/2020/07/09/europe-and-ai-leading-lagging-behind-or-carving-its-own-way-pub-82236.

249 Simon Pfeiffer and Randolf Carr, *Error 404–Trust Not Found: A European Survey on Digital (Dis)trust*, Munich Security Conference, Munich Security Brief 2, March 2021, https://doi.org/10.47342/REFQ1817.

250 See chapter 1.1 for more examples of privacy-preserving machine learning techniques.

251 Giovanni Sartor et al., "The Impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence: Study," European Parliamentary Research Service, June 2020, http://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf; and He Li, Lu Yu, and Wu He, "The Impact of GDPR on Global Technology Development," *Journal of Global Information Technology Management* 22, no. 1, January 2, 2019, 1–6, https://doi.org/10.1080/1097198X.2019.1569186.

252 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act).

253 European Commission, Proposal for a Regulation on European Data Governance (Data Governance Act), Article 5(3).

254 "Access Now's Position on the Data Governance Act," AccessNow (nonprofit organization), January 2021, https://www.accessnow.org/cms/assets/uploads/2021/03/Access-Nows-position-on-the-Data-Governance-Act.pdf; and Diego Naranjo, *EU Alphabet Soup of Digital Acts: DSA, DMA and DGA*, European Digital Rights (EDRi) Association, November 25, 2020, https://edri.org/our-work/eu-alphabet-soup-of-digital-acts-dsa-dma-and-dga/.

Geopolitical Considerations

Easier access to personal data gives countries such as China a data advantage in machine learning. Technology and security experts are concerned by the apparent systemic advantage political regimes like China's have in machine learning.[255] This includes the Chinese Communist Party's ability to access citizens' data without consent and its motivation to invest heavily in machine learning that enhances surveillance and social control, which has enabled the rapid development of artificial intelligence.[256] While China's Personal Information Protection Law (which comes into effect in November 2021) is modeled on GDPR,[257] experts note that China's approach to data-protection enforcement has focused more on investigating foreign companies and accessing their data than preventing the unauthorized transfer and use of personal data, which is widespread in China.[258]

Democracies such as EU member states face an existential challenge of protecting European fundamental rights while developing better performing AI systems based on larger and more diverse data sets.[259] Ethics and artificial intelligence expert Tim Hwang argues that "democracies may be in the unenviable position of having to compromise on core values like privacy to preserve their technological lead."[260] To overcome that challenge, Hwang suggests that democracies should invest in a set of technological solutions that would reduce or replace the reliance of machine learning systems on large data sets.[261]

Privacy-preserving machine learning could become the EU's competitive advantage. Differences in the emphasis on an individual's right to privacy and control of personal data is one of the key issues highlighting divergence in philosophical and regulatory approaches among key global actors such as the United States, China, and the EU.[262] The EU sees its emphasis on privacy, transparency, and fairness as its competitive advantage when technologies can increasingly undermine human rights.[263] The GDPR is a central element to this because it sets the EU apart from the United States, which still lacks any federal data-privacy legislation. [264] However, senior US officials have recently highlighted the importance of data security and privacy for the competitiveness of the United States, including specific mention of the potential of technologies like privacy preserving machine learning (PPML) by President Biden's National Security Advisor Jake Sullivan.[265] This may open opportunities for the EU to cooperate with the United States in accelerating the development and deployment of these technologies. PPML techniques such as federated learning could make it possible for EU-based companies to benefit from large and diverse data pools without compromising privacy (chapter 1.1 provides further details on PPML techniques).[266]

National Considerations

Citizens of EU member states expect government to ensure the protection of their personal data and some populations are less willing to share their data than others. Europeans share a high level of concern about the security of their personal data.[267] Approximately half of the respondents in a recent Munich Security Conference survey think that their national government is not acting aggressively enough on a range of security and privacy issues, such

255 Hwang, *Shaping the Terrain*.

256 Hwang, *Shaping the Terrain*.

257 Josh Horwitz, "China Passes New Personal Data Privacy Law, to Take Effect Nov. 1," Reuters, August 20, 2021, sec. China, https://www.reuters.com/world/china/china-passes-new-personal-data-privacy-law-take-effect-nov-1-2021-08-20/.

258 Lauren Maranto, "Who Benefits from China's Cybersecurity Laws?," *New Perspectives on Asia* (blog), Center for Strategic and International Studies (CSIS), June 25, 2020, https://www.csis.org/blogs/new-perspectives-asia/who-benefits-chinas-cybersecurity-laws.

259 I. Glenn Cohen et al., "The European Artificial Intelligence Strategy: Implications and Challenges for Digital Health," *The Lancet Digital Health* 2, no. 7 (July 2020): e376–79, https://doi.org/10.1016/S2589-7500(20)30112-6.

260 Hwang, *Shaping the Terrain*.

261 Hwang, *Shaping the Terrain*.

262 Matthew Goodman and Pearl Risberg, "Advancing Data Governance in the G7," CSIS, February 2, 2021, https://www.csis.org/analysis/advancing-data-governance-g7.

263 Brattberg, Csernatoni, and Rugova, *Europe and AI*.

264 Rachel F. Fefer and Kristin Archick, "EU Data Protection Rules and U.S. Implications," *In Focus* (brief), Congressional Research Service, IF10896, July 17, 2020.

265 The White House, "Remarks by National Security Advisor Jake Sullivan at the National Security Commission on Artificial Intelligence Global Emerging Technology Summit," White House (Briefing Room website), July 13, 2021, https://www.whitehouse.gov/nsc/briefing-room/2021/07/13/remarks-by-national-security-advisor-jake-sullivan-at-the-national-security-commission-on-artificial-intelligence-global-emerging-technology-summit/.

266 Federated learning is a technique that trains an AI algorithm across decentralized devices or servers holding data samples without exchanging those samples, enabling multiple parties to build a common machine learning model without sharing data liberally. More technical details on privacy-preserving machine learning is out of the scope of this report, but the following resources provide a good overview: Hwang, *Shaping the Terrain*; Jason Mancuso, "Privacy-Preserving Machine Learning 2019: A Year in Review," Medium, Cape Privacy (formerly Dropout Labs), January 10, 2020, https://medium.com/dropoutlabs/privacy-preserving-machine-learning-2019-a-year-in-review-123733e61705; Patricia Thaine, "Perfectly Privacy-Preserving AI," Medium, January 2, 2020, https://towardsdatascience.com/perfectly-privacy-preserving-ai-c14698f322f5; and Kyle Wiggers, "AI Has a Privacy Problem, but These Techniques Could Fix It," VentureBeat, December 21, 2019, https://venturebeat.com/2019/12/21/ai-has-a-privacy-problem-but-these-techniques-could-fix-it/.

267 Pfeiffer and Carr, *Error 404–Trust Not Found*.

Figure 4: European Trust in Safety of Personal Data, 2021 (percent)[1]

(Data: Kekst CNC, commissioned by the Munich Security Conference, General population in France, Germany, Italy, Poland, Sweden and the UK; excluding "don't know" and neutral responses).

1 Pfeiffer and Carr, *Error 404–Trust Not Found*.

as how technology companies protect data.[268] Only about one-quarter of respondents had confidence in companies from other EU member states protecting their data, less than one-fifth had confidence in US companies and one-tenth in Chinese companies (see figure 4).[269] "A 2019 survey showed that on average people in Finland were more likely to share their personal data to improve public services like medical research and care (66 percent), people in Germany and France were less likely (44 percent and 40 percent), and people in Poland unlikely (30 percent).[270] The European Commission needs to consider how to close this gap when developing policies to enable machine learning across the EU.

However, some member states are also concerned that overregulation has the potential to stifle innovation. In October 2020, Denmark drafted a document co-signed by France, Finland, and Poland alongside eleven other EU member states advocating to "ensure that the European Commission sees innovative artificial intelligence as equally important as managing risks."[271] Therefore, the commission needs to carefully balance measures that will increase trust in privacy protection as well as enhance the competitiveness of European businesses.

Business Considerations

Large volumes of training data, which often include personal data, are essential for the development of machine learning systems. For example, to launch safe autonomous vehicles, companies rely on vast amounts of training data to teach and test vehicles on how to navigate various complex driving situations safely.[272] Training data availability is therefore at the core of the advancement

268 Pfeiffer and Carr, *Error 404–Trust Not Found*.

269 Pfeiffer and Carr, *Error 404–Trust Not Found*.

270 "Attitudes towards the Impact of Digitalisation on Daily Lives," Eurobarometer (polling instrument), European Union (website), March 2020, https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2228.

271 Janosch Delcker, "14 EU Countries Urge Brussels to Go Easy on AI Laws," *Politico Pro* (subscription platform), August 10, 2020.

272 Sandra Wachter, "Autonomous Vehicles and Risks to Privacy," Video Remarks of Oxford Internet Institute (OII) Expert, Posted by Harvard's Berkman Klein Center for Internet & Society, 2018, https://www.youtube.com/watch?v=WZfV22YHdlU. Wachter is associate professor and senior research fellow in law and ethics of AI, Big Data, and robotics at University of Oxford's OII and a faculty associate at the Berkman Klein Center.

of autonomous vehicles companies' programs.[273] Since those vehicles interact with many stakeholders, such as car owners and passengers, pedestrians, or other drivers, they also generate unprecedented volumes of personal data of which GDPR restricts usage.[274]

Businesses can face high GDPR compliance costs or legal uncertainty that can deter the development of machine learning. Early evidence from the GDPR suggests that for some businesses, compliance costs can be significant and the rules complex to implement.[275] For example, the GDPR requires consent from the subject before collecting or using personal data, but this becomes extremely complex to implement in cases like autonomous vehicles where there are a high number of stakeholders and large volumes and categories of data being collected.[276] Potential steep fines also can make businesses hesitant about using and sharing personal data, even if adequate protections are in place. The European Commission should introduce additional measures to diminish the costs of compliance and legal uncertainty while enhancing data-protection enforcement.

Recommendations for the European Commission

A. Coordinate with the European Data Protection Board and national data protection authorities to provide more detailed guidance and assistance on implementing the GDPR to machine learning applications to enhance compliance and mitigate business costs.

This could include work on establishing potential use cases, guidelines, and technical standards,[277] which would decrease the costs of compliance, risks from legal uncertainty, and allay concerns that stem from the possibility of high fines.[278]

B. Enhance investment in research in privacy-preserving machine learning techniques that can adequately protect personal data while enabling machine learning development, including in partnership with like-minded countries such as the United States.

The development of techniques that can viably replace the need to compile large pools of personal data for machine learning could enable the EU to better innovate using machine learning while upholding liberal democratic values.[279] The commission should consider enhancing its cooperation with the United States in developing these techniques. For example, the commission could propose incorporating them in one of the working groups of the EU-US Trade and Technology Council, which was launched in June 2021. [280] It could also expand the scientific cooperation between the European Commission's Joint Research Centre (JRC) and the US National Institute of Standards and Technology (NIST).

3.2 Mitigate data-driven discrimination

Machine learning systems leverage training data sets to observe patterns and automate prediction and decision-making. However, without safeguards, machine learning may reproduce and scale existing discriminatory systems as well as biases that can violate human rights.[281] While there are several ways this can happen, data is one of the key contributing factors leading to discriminatory machine learning predictions and outcomes.[282] Currently, there are gaps in the EU's legal and technical measures available to address the issue, which could lead to societal and reputational costs for the EU, member states, and organizations. **The European Commission needs to address the gaps in current rules and guidelines on machine learning-enabled discrimination while avoiding negatively impacting the competitiveness of EU businesses and the EU's potential to shape global AI standards.**

273 Deloitte, *Connected and Autonomous Vehicles in Ontario: Implications for Data Access, Ownership, Privacy and Security*, n.d., https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/consulting/ca-consulting-CVAV-Research-Final-Data-Privacy-Security-Report-20180412-AODA.pdf; and Kelsey Piper, "It's 2020. Where Are Our Self-Driving Cars?," *Vox*, February 28, 2020, https://www.vox.com/future-perfect/2020/2/14/21063487/self-driving-cars-autonomous-vehicles-waymo-cruise-uber.

274 Deloitte, *Connected and Autonomous Vehicles in Ontario*.

275 Alex Marthews and Catherine Tucker, *Privacy Policy and Competition*, Brookings Institution, December 2019, 10, https://www.brookings.edu/wp-content/uploads/2019/12/ES-12.04.19-Marthews-Tucker.pdf.

276 Stephan Appt, "Legal Challenges of Data-Driven Mobility," Pinsent Masons (law firm's website), November 4, 2020, https://www.pinsentmasons.com/out-law/analysis/legal-challenges-data-driven-mobility.

277 Andrew Imbrie et al., *Agile Alliances: How the United States and Its Allies Can Deliver a Democratic Way of AI,* Center for Security and Emerging Technology, Georgetown University's Walsh School of Foreign Service, February 2020, https://cset.georgetown.edu/research/agile-alliances/.

278 Sartor et al., "The Impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence."

279 Hwang, *Shaping the Terrain*.

280 "EU-US Launch Trade and Technology Coucil to Lead Values-Based Global Digital Transformation," Press Release, European Commission (website), June 15, 2021, https://ec.europa.eu/commission/presscorner/detail/en/IP_21_2990.

281 Andrea Renda, *Artificial Intelligence*, CEPS, February 15, 2019, https://www.ceps.eu/ceps-publications/artificial-intelligence-ethics-governance-and-policy-challenges/; and French Data Protection Authority (CNIL), *How Can Humans Keep the Upper Hand? Report on the Ethical Matters Raised by Algorithms and Artificial Intelligence*," Commission Nationale Informatiques et Libertés (CNIL), December 2017, https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf.

282 Smith and Rustagi, *Mitigating Bias in Artificial Intelligence*.

To address this, the European Commission should clarify the distribution of responsibilities between the enforcement authorities of existing rules and newly emerging authorities proposed in upcoming AI regulation to avoid overlaps or loopholes in implementation. The commission should provide sufficient funding across the member states for data protection authorities and other relevant bodies that enforce the existing rules to minimize gaps in implementation. It should also increase cooperation on reducing discrimination in machine learning with the United States.

EU Policy Context

Existing data protection and nondiscrimination legislation provides a robust framework but fails to fully capture the problems emerging from discrimination enabled by machine learning. For example, while the GDPR helps mitigate unfair and illegal discrimination based on personal data, it does not address those risks in nonpersonal data.[283] The GDPR introduced a mandatory Data Protection Impact Assessment (DPIA) to examine the likely impact of data processing on individuals' rights and fundamental freedoms, but it only applies in certain "high-risk" cases, for example when organizations conduct profiling of individuals.[284] The currently available legal tools to deter discrimination are also not always easily transferrable to machine learning, for example, because discriminatory outputs can be hard to detect.[285] The European Commission has also provided voluntary "Trustworthy AI assessment tools" to fill this gap.[286] However, there remains a lack of oversight mechanisms in place to incentivize compliance.[287]

The new EU AI regulation proposed by the European Commission in April 2021 outlines data-governance and -management requirements to protect fundamental rights but lacks clarity on implementation.[288] The proposal aims to complement existing nondiscrimination laws with additional requirements to minimize the risk of algorithmic discrimination, especially in relation to the design and the quality of data sets used for the development of AI systems. It includes requirements for training, validation, and testing data sets to be "sufficiently relevant, representative and free of errors, and complete."[289] Article 10 specifies the "quality criteria" for data sets used for the development of "high-risk" AI systems, such as remote biometric identification, including design choices, data collection, processing (e.g., labeling and cleaning), assessment of the data sets, examination of possible biases, or identification of any possible data gaps or shortcomings.[290] While providing data requirements to mitigate potential for discrimination is a step in the right direction, the requirements appear to lack specificity and clarity on how they would be implemented, potentially relying on self-assessment of AI system providers, whose interest primarily rests with ensuring compliance and who may lack necessary expertise.[291]

Furthermore, the proposed measures to reduce discrimination may have limited effect and will require significant investment in human resources and expertise by member states as well as coordination among European regulatory bodies. Technology and law experts Kenneth Propp and Mark MacCarthy argue that the proposal's specific mechanism designed to monitor, detect, and correct bias in high-risk AI systems may have limited effect on addressing discrimination given disparities in required impact assessments. They also highlight the limited impact of the proposed documentation requirements of AI systems since the documentation does not need to be provided to those potentially affected by the systems, only to regulators upon request.[292] Furthermore, the new rules will require complex

283 Frederik Zuiderveen Borgesius, *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*, Council of Europe, Directorate General of Democracy, 2018, https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73; and Institut Montaigne, *Algorithms: Please Mind the Bias!*, March 2020, https://www.institutmontaigne.org/ressources/pdfs/publications/algorithms-please-mind-bias.pdf.

284 European Union Agency for Fundamental Rights, *Getting the Future Right*; and "When Is a Data Protection Impact Assessment (DPIA) Required?," European Commission (website), accessed August 16, 2021, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/when-data-protection-impact-assessment-dpia-required_en.

285 Sandra Wachter, "Why AI Fairness Cannot Be Automated," Presentation at Oxford Internet Institute, University of Oxford, July 2020, https://www.youtube.com/watch?v=p8MCaj68Pns; and Sara Spinks, "Algorithmic Bias within Online Behavioural Advertising Means Public Could Be Missing Out, says Associate Professor Sandra Wachter," Oxford Internet Institute (website), November 2019, https://www.oii.ox.ac.uk/blog/algorithmic-bias-within-online-behavioural-advertising-means-public-could-be-missing-out-says-associate-professor-sandra-wachter/.

286 The European Commission set up a High-Level Expert Group on Artificial Intelligence (AI HLEG) that presented in July 2020 its assessment list for trustworthy AI, a self-assessment tool to "ensure that users benefit from AI without being exposed to unnecessary risks." See "Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment," European Commission (website), https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment; and Access Now, *Europe's Approach to Artificial Intelligence: How AI Strategy Is Evolving*," Access Now (website), December 2020, https://www.accessnow.org/cms/assets/uploads/2020/12/Europes-approach-to-AI-strategy-is-evolving.pdf.

287 Access Now, "Europe's Approach to Artificial Intelligence."

288 European Commission, "Proposal for a Regulation Laying down Harmonised Rules, pp. 48-50 and 53.

289 European Commission, "Proposal for a Regulation Laying down Harmonised Rules"; and Kenneth Propp and Mark MacCarthy, "Machines Learn That Brussels Writes the Rules: The EU's New AI Regulation," Brookings (website), May 4, 2021, First Published in *Lawfare* (blog), https://www.brookings.edu/blog/techtank/2021/05/04/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/.

290 European Commission, "Artificial Intelligence Act."

291 Sarah Chander and Ella Jakubowska, "EU's AI Law Needs Major Changes to Prevent Discrimination and Mass Surveillance," EDRi (website), May 2021, https://edri.org/our-work/eus-ai-law-needs-major-changes-to-prevent-discrimination-and-mass-surveillance/.

292 Propp and MacCarthy, "Machines Learn That Brussels Writes the Rules."

implementation. Member states would need to create new AI regulatory authorities to implement and ensure compliance of these rules, including coordination with other regulatory bodies with intersecting responsibilities.

Geopolitical Considerations

China is increasingly influential in exporting and shaping AI surveillance technologies and norms globally, which creates risks for upholding human rights, including the right to nondiscrimination. Experts are concerned that China's deployment of machine learning includes leveraging surveillance technologies to target specific groups.[293] For example, Chinese data scientists have trained a facial recognition algorithm on ethnographic data to recognize the faces of Uighur people, a predominantly Muslim ethnic group in China.[294] The Chinese government is alleged to be leveraging this technology to track people in the Uighur community, according to a news report citing five people with direct knowledge of the systems, who were granted anonymity by the *New York Times*.[295] In reaction to these developments, a resolution adopted by the European Parliament in December 2020 stated: "The European Parliament strongly condemns the extensive use of digital surveillance technologies to monitor and control the population in Xinjiang," and expressed concern that China is exporting surveillance technologies to authoritarian regimes around the world.[296] An example of these exports is Chinese company Huawei Technologies, with research showing that the company is the leading vendor of AI surveillance technology worldwide and that it is shaping standards by embedding them in products, such as facial recognition systems, deployed across at least fifty countries.[297]

Western democracies can set international standards that limit data-driven discrimination in machine learning, but they lack an aligned approach. Despite their shared values, the United States and the EU are not closely aligned in many aspects of governing technology. For example, the White House Office of Science and Technology issued a statement in January 2020 arguing that "Europe and our allies should avoid heavy-handed innovation-killing models." The statement continued by saying that to shape the evolution of technology and "counter authoritarian uses of AI," the United States and its partners need to remain the global hubs of innovation.[298] There is some acknowledgment of this view in Europe—in October 2020, twenty-six EU member states agreed that the EU approach needs to provide "sufficient flexibility in research and development of these [AI] systems."[299] However, EU member states put a heavy emphasis on the need for high ethical standards.[300] To effectively influence international standards, the European Commission should intensify cooperation with US authorities to seek areas of cooperation, and the newly established EU-US Trade and Technology Council (TTC) is therefore a step in the right direction.[301]

National Considerations

There is a growing awareness of the risks of discrimination in AI systems, causing more concerns in some European countries than others. A 2019 survey showed that citizens of France (46 percent) and Germany (43 percent) are more likely to be concerned that AI could lead to discrimination than those of Finland (35 percent) and Poland (28 percent).[302] Additionally, EU member states provide varying levels of resources to institutions investigating data-protection and antidiscrimination cases (see figure 5). For example, Germany significantly leads both in terms of funding and the number of tech specialists that work for its Data Protection Authorities.[303] The European Commission will have to navigate and address differences in priorities and funding for protections against data-driven discrimination that could undermine a unified European approach.

293 Andrea Kendall-Taylor, Erica Frantz, and Joseph Wright, "The Digital Dictators," *Foreign Affairs*, February 2, 2021, https://www.foreignaffairs.com/articles/china/2020-02-06/digital-dictators; and Erol Yayboke and Samuel Brannen, "Promote and Build: A Strategic Approach to Digital Authoritarianism," CSIS Brief, October 2020, https://www.csis.org/analysis/promote-and-build-strategic-approach-digital-authoritarianism.

294 Cunrui Wang et al., "Facial Feature Discovery for Ethnicity Recognition," *WIREs Data Mining and Knowledge Discovery* 9, no. 1 (August 2018): e1278, https://doi.org/10.1002/widm.1278.

295 Paul Mozur, "One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority," *New York Times*, April 2019, https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html.

296 Forced Labour and the Situation of the Uighurs in the Xinjiang Uighur Autonomous Region, European Parliament Res., P9_TA(2020)0375 (December 17, 2020), https://www.europarl.europa.eu/doceo/document/TA-9-2020-0375_EN.html.

297 Steven Feldstein, *The Global Expansion of AI Surveillance*, Carnegie Endowment for International Peace, accessed April 20, 2021, https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847.

298 Samuel Stolton, "Avoid Heavy AI Regulation, White House Tells EU," Euractiv, January 2020, https://www.euractiv.com/section/digital/news/avoid-heavy-ai-regulation-white-house-tells-eu/.

299 "The Charter of Fundamental Rights in the Context of AI and Digital Change," Note from the Presidency to Delegations, 11481/20, Council of the European Union, October 21, 2020, https://www.consilium.europa.eu/media/46496/st11481-en20.pdf.

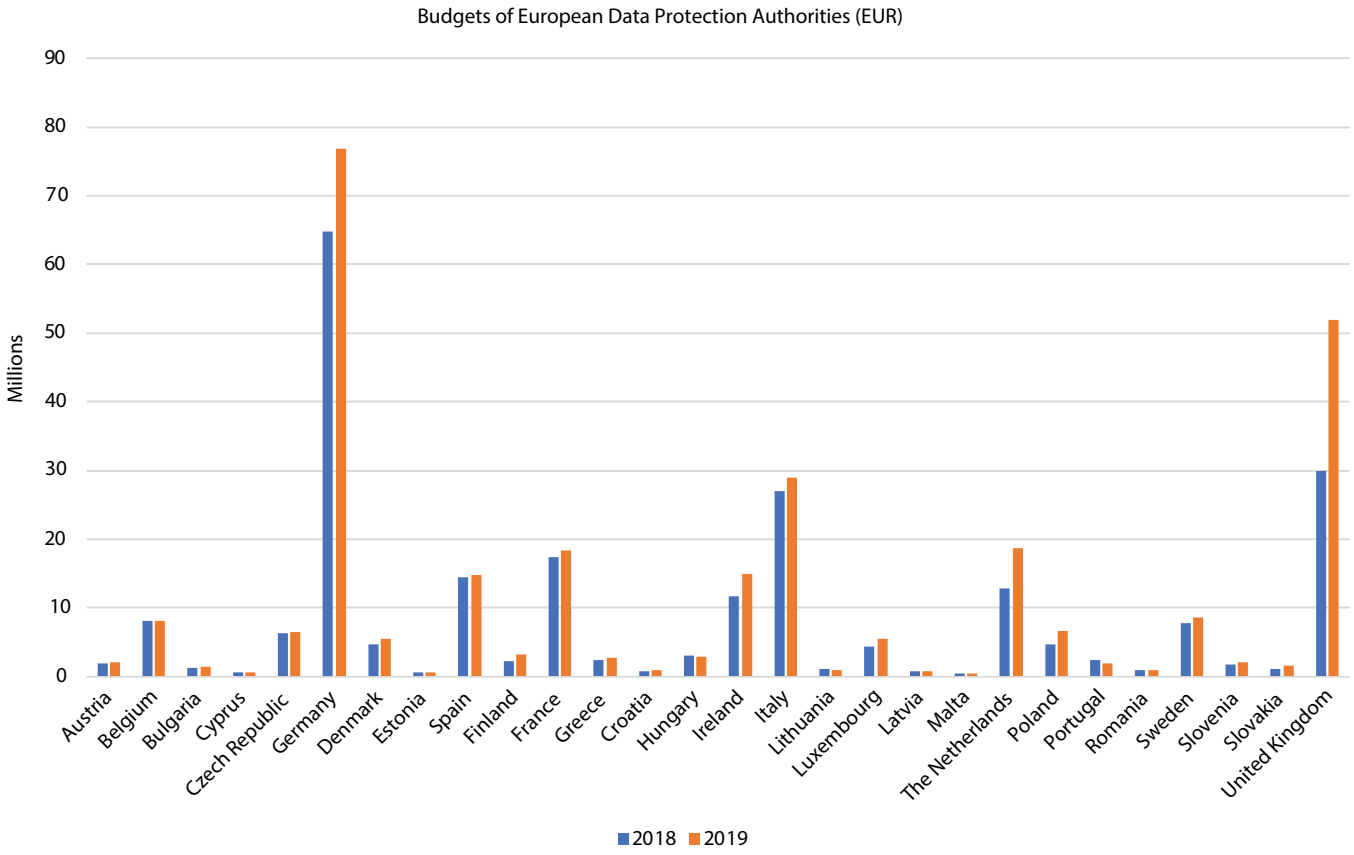
300 "The Charter of Fundamental Rights in the Context of AI and Digital Change."

301 European Commission, "EU-US Launch Trade and Technology Council."

302 European Commission, "Standard Eurobarometer 92," December 2019, https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/STANDARD/surveyKy/2255; and European Union Agency for Fundamental Rights, *Getting the Future Right*.

303 Estelle Massé, *Two Years under the EU GDPR: An Implementation Progress Report*, Access Now, May 2020, https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf.

Figure 5: Differences in the Budgets of European Data Protection Authorities in 2018 and 2019[1]



1 Massé, *Two Years under the EU GDPR: An Implementation Progress Report.*

Businesses Considerations

Discrimination caused by machine learning can harm customers or product quality, posing significant risks for businesses. Biased systems result in inaccurate predictions and outputs that can be discriminatory. They can unfairly allocate opportunities or negatively impact people's well-being.[304] Biased systems can consequently lead to infringements of human rights and consumer protections, resulting in costly system overhauls and liability disputes for damages caused.[305] They also can negatively impact a business's reputation, harming consumer trust and future market opportunities.[306] For example, in 2019, a Microsoft filing to the US Securities and Exchange Commission flagged the potential for reputational harm due to "flawed" AI algorithms or data sets with "biased information."[307]

Businesses are aware of the importance of responsible AI but are unsure how to address the issue in practice. For example, a 2021 survey by PwC found that over two-thirds of US businesses were not taking action to reduce bias from AI.[308] A 2020 report by the UK Centre for Data Ethics and Innovation concluded that many organizations are unsure of how to address bias in practice, including choosing the right mitigation methods and embedding them into the operations of the organization.[309]

304 Smith and Rustagi, *Mitigating Bias in Artificial Intelligence*.

305 Barbero et al., *Study on Emerging Issues of Data Ownership*.

306 "Understanding Algorithmic Bias and How to Build Trust in AI," PwC (website), January 18, 2021, https://www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions-2021/algorithmic-bias-and-trust-in-ai.html; and Smith and Rustagi, *Mitigating Bias in Artificial Intelligence*."

307 Microsoft Corporation, Form 10-Q for the Quarter Ended December 31, 2018, US Securities and Exchange Commission, File Number 001-37845, January 2019, https://www.sec.gov/Archives/edgar/data/789019/000156459019001392/msft-10q_20181231.htm.

308 PwC, "Algorithmic Bias and Trust in AI."

309 Centre for Data Ethics and Innovation, "Independent Report: Review into Bias in Algorithmic Decision-Making," UK Government (website), November 2020, https://www.gov.uk/government/publications/cdei-publishes-review-into-bias-in-algorithmic-decision-making/main-report-cdei-review-into-bias-in-algorithmic-decision-making.

While business can benefit from clear guidelines, they are wary of the economic cost of strict regulations and their impact on innovation.[310] In 2019, the executive vice-president of the European Commission, Margrethe Vestager, said that the ethics guidelines established by the EU's High-Level Expert Group saw "a very high number of businesses sign up," showing a demand for more guidance.[311] However, similar to regulatory measures to protect data privacy (see Chapter 3.1), while businesses can benefit from guidance, overly strict rules aimed at limiting data-driven discrimination could pose compliance costs that prohibit the development of new products and services, particularly for smaller businesses.[312] The European Commission needs to introduce rules that would increase companies' uptake of the ethical principles without unnecessarily stifling innovation and hampering start-ups.[313]

Recommendations for the European Commission

This section addresses the issue of discrimination specifically in the context of data and machine learning and offers the following recommendations.[314]

A. Work with the European Parliament and member states on providing more specific requirements for data-management standards to mitigate risks for discrimination.

This could include adding additional mechanisms that would diversify the reliance on self-assessment by system providers as well as requiring more transparency and broader public access to the bias assessment documentation of AI systems. It could also include ensuring that the proposal addresses discrimination that can stem from data sets that might be of high quality but reflect existing societal biases, and therefore represent risk for the AI systems to embed and amplify discrimination.

B. Provide additional financial and human resources to national data-protection authorities and other relevant bodies that investigate antidiscrimination violations, and carefully define and coordinate the establishment of new enforcement authorities with the existing ones.

Since the implementation of the upcoming EU AI regulation will rely on supervisory authorities of member states, this would enable the relevant authorities to better enforce existing laws by addressing disparities in the budgets of those authorities across the EU. Additionally, the newly established enforcement authorities, such as the EU Artificial Intelligence Board, and the national authorities need to be well integrated with the existing data-protection authorities. The European Commission should carefully define the roles and responsibilities of the newly established authorities to prevent potential overlaps or loopholes in the implementation of the regulation and to tap into existing resources and expertise.

C. Increase transatlantic cooperation to close the gap between the US and the EU approaches to mitigating risks of bias and discrimination enabled by data in machine learning.

While the United States is likely to maintain a decentralized approach to regulating AI, the EU-US TTC, launched in June 2021, provides an opportunity to narrow the gaps in the EU and US regulatory approaches. The TTC should consider incorporating the issues of bias and discrimination enabled by data in machine learning in the agenda of two of its working groups (the data governance and technology platforms and the misuse of technology threatening security and human rights group).[315] The issue of addressing the potential for replicating or scaling discrimination could also be addressed in the Transatlantic AI agreement that the European Commission and the high representative of the Union for foreign affairs and security policy proposed to start working on in a joint December 2020 communication, "A New EU-US Agenda for Global Change."[316]

310 Madhumita Murgia, "The Four Problems with Europe's Vision of AI," *Financial* Times, February 2020, https://www.ft.com/content/6759046a-57bf-11ea-a528-dd0f971febbc.

311 *Europe Fit for the Digital Age: Hearing of Margrethe Vestager, Executive Vice-President-Designate of the European Commission*, Committees on Industry, Research, and Energy, on the Internal Market and Consumer Protection, and on Economic and Monetary Affairs, European Parliament (October 8, 2019), https://www.europarl.europa.eu/resources/library/media/20191009RES63801/20191009RES63801.pdf. See also: "Ethics Guidelines for Trustworthy AI," European Commission (website), April 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

312 Digital Curation Centre, University of Edinburgh, *The Role of Data in AI*, Prepared for the Data Governance Working Group, Global Partnership of AI, November 2020, https://gpai.ai/projects/data-governance/role-of-data-in-ai.pdf.

313 Brattberg, Csernatoni, and Rugova, *Europe and AI*.

314 Given this focus, this section omits other important factors and measures that are key to addressing discrimination in artificial intelligence systems in their full complexity.

315 European Commission, "EU-US Launch Trade and Technology Council."

316 European Commission and the High Representative of the Union for Foreign Affairs and Security Policy, "A New EU-US Agenda for Global Change," Joint Communication to the European Parliament, the European Council, and the Council, December 2, 2020, https://ec.europa.eu/info/sites/default/files/joint-communication-eu-us-agenda_en.pdf.

Conclusion

Machine learning will have a transformative effect on industries and societies. The rules on who can access data and how it can be used for machine learning will have significant bearing on an economy's ability to take advantage of the technology. To compete with Chinese and US industries, the EU will need data policies that spur innovation through machine learning. However, to mitigate the negative consequences on fundamental rights and European autonomy, it will equally need policies that address these risks.

This report provides European decision makers with a set of recommended actions designed to unlock the potential of machine learning, while mitigating the risks (see summary of recommendations). These recommendations are based on a close examination of the geopolitical considerations of the EU, domestic considerations of EU member states, and commercial considerations of businesses. Using this framework, the following are the key policy objectives that our recommendations intend to achieve. They represent what we believe the European Commission should pursue in relation to data rules for machine learning.

First, increase access to public-sector data by:

- addressing fragmented open data rules and procedures across Europe to reduce the administrative and legal costs for businesses seeking to reuse public-sector data
- allowing access to sensitive data with appropriate privacy preserving techniques and protections in place to accelerate public and private research and development

Second, harmonize data standards for interoperability by:

- facilitating the development of tools and protocols that enable the seamless exchange and translation of differently structured data across systems and organizations in Europe.

Third, diversify data storage and processing by:

- ensuring coordination of proposed and existing cloud initiatives and minimizing compliance costs of new regulatory measures
- encouraging additional investment in data infrastructure services to better ensure competitive cloud offerings for EU-based businesses
- refraining from protectionist data processing and storage measures that have a more negative impact on smaller and more trade-dependent EU member states

Fourth, simplify data transfers by:

- developing new cross-border data mechanisms that ensure data protection while enabling EU-based businesses to participate in global value chains
- reducing legal ambiguity in EU legislation that increases risks and costs for businesses sharing data

Fifth, foster inclusive data ecosystems by:

- adopting sector-specific data portability rights and introducing rules on when businesses should allow access to proprietary data sets.

Sixth, enhance privacy and trust by:

- clarifying guidance on applying the GDPR to machine learning techniques to reduce the costs of compliance and legal uncertainty while enhancing data protection
- investing in research in privacy-preserving machine learning techniques and enhancing cooperation with like-minded countries like the United States

Seventh, mitigate data-driven discrimination by:

- clarifying responsibilities of existing and proposed authorities to enforce non-discrimination laws to avoid overlaps or loopholes in the implementation
- ensuring national data-protection authorities and other enforcement agencies are adequately funded
- increasing transatlantic cooperation on mitigating the risks of bias and discrimination enabled by data in machine learning

The EU is falling behind China and the United States in its ability to harness machine learning. This has the potential to weaken the ability of EU-based businesses to compete in a global market, reinforce the EU's dependencies on foreign digital technologies, and undermine its ability to shape global rules at the intersection of technology and society. However, if the EU manages to strike the right

balance in its regulatory approach by unlocking the opportunities for businesses while preserving robust protections for fundamental rights, it could put the EU on a more sustainable footing to leverage machine learning technologies. If the commission succeeds, not only will it enhance European economic prosperity and deliver tools that can help address society's most pressing challenges, Europe will also serve as a model for other countries to follow in achieving innovative policies that benefit and deliver value to society.

European Union Commissioner for Competition Margrethe Vestager speaks at a Friends of Europe event. *Source:* Flickr/Friends of Europe (https://www.flickr.com/photos/friendsofeurope/17223939225)

Acknowledgments & Author Bios

Blanka Soulava is a Master in Public Policy Candidate at the Harvard Kennedy School. Previously, she worked as a Trainee at the Office of the Vice-President of the European Parliament and as a campaign assistant and advisor during the 2018 Czech presidential elections. Blanka's policy interests include digital and tech policy and negotiation. She has a bachelor's degree in International Area Studies from Charles University in Prague, a diploma in Economics from the University of London, and studied Chinese at Fudan University in Shanghai.

Hamish Cameron was a Graham T. Allison, Jr. Student Fellow at the Harvard Belfer Center and completed a Master in Public Policy at the Harvard Kennedy School in 2019-21. He previously served as a diplomat in Beijing with the Australian foreign service. He has a background in strategic economic issues including digital governance, energy, and climate change, and he developed the Australian Government's 2021 Services Exports Action Plan. He has a Bachelor of Arts and a Bachelor of Science from the University of Melbourne.

Victoria Ying is currently a Master of Public Policy candidate at Harvard Kennedy School (HKS). At HKS, her academic interests include machine learning, privacy and security, and misinformation and disinformation campaign. Outside of HKS, she is building organizational partnerships to advance tech-enabled negotiation solutions.

This report would not have been possible without the generous help and support of Harvard Kennedy School Professors Bruce Schneier and Thomas Patterson, as well as Trey Herr and Safa Shahwan Edwards at the Atlantic Council's Cyber Statecraft Initiative. We would also like to thank our families and loved ones who supported us in our work throughout the challenging pandemic year.

Additionally, this report benefitted significantly from the time and insights shared by experts from across government, industry, academia and civil society. For this, we would like to acknowledge the following people: Alex Cooke, Andrea Renda, Antonio Capobianco, Christie Lawrence, Daniel Leufer, Eline Chivot, Grace Abuhamad, Gretchen Greene, Henrique Choer Maraes, Iga Kozlowska, Jaana Sinipuro, James Mancini, Jörn Fleck, Kenneth Propp, Kurt John, Luc Nicolas, Marc Lange, Markus Kalliola, Mayank Agrawal, Nicolai van Gorp, Raluca Csernatoni, Valentina Pavel, Werner Stengg, Yushi Wang.

Atlantic Council

Board of Directors

CHAIRMAN
*John F.W. Rogers

EXECUTIVE CHAIRMAN EMERITUS
*James L. Jones

PRESIDENT AND CEO
*Frederick Kempe

EXECUTIVE VICE CHAIRS
*Adrienne Arsht
*Stephen J. Hadley

VICE CHAIRS
*Robert J. Abernethy
*Richard W. Edelman
*C. Boyden Gray
*Alexander V. Mirtchev
*John J. Studzinski

TREASURER
*George Lund

DIRECTORS
Stéphane Abrial
Todd Achilles
*Peter Ackerman
Timothy D. Adams
*Michael Andersson
David D. Aufhauser
Barbara Barrett
Colleen Bell
Stephen Biegun
*Rafic A. Bizri
*Linden P. Blue
Adam Boehler
Philip M. Breedlove
Myron Brilliant
*Esther Brimmer
R. Nicholas Burns
*Richard R. Burt
Teresa Carlson
James E. Cartwright
John E. Chapoton
Ahmed Charai
Melanie Chen
Michael Chertoff
*George Chopivsky
Wesley K. Clark
*Helima Croft
Ralph D. Crosby, Jr.
*Ankit N. Desai
Dario Deste
*Paula J. Dobriansky
Joseph F. Dunford, Jr.
Thomas J. Egan, Jr.
Stuart E. Eizenstat
Thomas R. Eldridge
Mark T. Esper
*Alan H. Fleischmann
Jendayi E. Frazer
Courtney Geduldig
Meg Gentle
Thomas H. Glocer
John B. Goodman
*Sherri W. Goodman
Murathan Günal
Amir A. Handjani
Frank Haun
Michael V. Hayden
Tim Holt
*Karl V. Hopkins
Andrew Hove
Mary L. Howell
Ian Ihnatowycz
Mark Isakowitz
Wolfgang F. Ischinger
Deborah Lee James
Joia M. Johnson
*Maria Pica Karp
Andre Kelleners
Henry A. Kissinger
*C. Jeffrey Knittel
Franklin D. Kramer
Laura Lane
Jan M. Lodal
Douglas Lute
Jane Holl Lute
William J. Lynn
Mark Machin
Mian M. Mansha
Marco Margheri
Michael Margolis
Chris Marlin
William Marron
Gerardo Mato
Timothy McBride
Erin McGrain
John M. McHugh
Eric D.K. Melby
*Judith A. Miller
Dariusz Mioduski
*Michael J. Morell
*Richard Morningstar
Georgette Mosbacher
Dambisa F. Moyo
Virginia A. Mulberger
Mary Claire Murphy
Edward J. Newberry
Thomas R. Nides
Franco Nuschese
Joseph S. Nye
Ahmet M. Ören
Sally A. Painter
Ana I. Palacio
*Kostas Pantazopoulos
Alan Pellegrini
David H. Petraeus
W. DeVier Pierson
Lisa Pollina
Daniel B. Poneman
*Dina H. Powell McCormick
Ashraf Qazi
Robert Rangel
Thomas J. Ridge
Gary Rieschel
Lawrence Di Rita
Michael J. Rogers
Charles O. Rossotti
Harry Sachinis
C. Michael Scaparrotti
Ivan A. Schlager
Rajiv Shah
Gregg Sherrill
Ali Jehangir Siddiqui
Kris Singh
Walter Slocombe
Christopher Smith
Clifford M. Sobel
James G. Stavridis
Michael S. Steele
Richard J.A. Steele
Mary Streett
*Frances M. Townsend
Clyde C. Tuggle
Melanne Verveer
Charles F. Wald
Michael F. Walsh
Ronald Weiser
Olin Wethington
Maciej Witucki
Neal S. Wolin
*Jenny Wood
Guang Yang
Mary C. Yates
Dov S. Zakheim

HONORARY DIRECTORS
James A. Baker, III
Ashton B. Carter
Robert M. Gates
James N. Mattis
Michael G. Mullen
Leon E. Panetta
William J. Perry
Condoleezza Rice
Horst Teltschik
William H. Webster

**Executive Committee Members*

List as of October 20, 2021

The Atlantic Council is a nonpartisan organization that promotes constructive US leadership and engagement in international affairs based on the central role of the Atlantic community in meeting today's global challenges.

Atlantic Council

1030 15th Street, NW, 12th Floor, Washington, DC 20005

(202) 463-7226, www.AtlanticCouncil.org