

## Comments for the Record



# Request for Information (RFI) Related to NIST's Assignments Under Sections 4.1, 4.5, and 11 of the Executive Order Concerning Artificial Intelligence

JOSHUA LEVINE, DANNY DOHERTY | FEBRUARY 2, 2024

## Introduction and Summary

Establishing guidelines and best practices for the development, deployment, and evaluation of artificial intelligence (AI) technologies is important to promote AI diffusion and adoption throughout the economy. The National Institute for Standards and Technology's (NIST) prior experience developing the AI Risk Management Framework (RMF) and the Cybersecurity Framework make it an ideal candidate to evaluate and develop guidelines for the promotion of safe, secure, and trustworthy AI systems.<sup>[1]</sup>

As NIST considers developing such guidelines, it should prioritize flexibility for technical solutions related to content provenance and watermarking, base guidance on existing red-teaming strategies and ensure new strategies can be developed and incorporated, consider how existing guidance such as the RMF and Cybersecurity Framework can supplement new proposals, and identify areas where guidelines can align with international frameworks related to AI-powered technologies. Throughout these areas, and as with the RMF and Cybersecurity Framework, this guidance should be iterative and evolutionary to respond to changing conditions of AI technologies, uses, and governance mechanisms.

## Developing Guidelines, Standards, and Best Practices for AI Safety and Security

### Economic and Security Implications of Watermarking, Provenance Tracking, and Other Tools

The RFI asks about the economic and security implications that watermarking, provenance tracking, and other authentication tools could have on generative AI. As generative AI tools proliferate and improve, people increasingly ask the question: Is this real, or is it AI? And while deceptive content is nothing new<sup>[2]</sup>, generative AI and other AI-powered tools drastically lower the cost to create forgeries or mimic the style or appearance of work done previously by humans.<sup>[3]</sup> As NIST crafts guidelines to promote the diffusion of existing tools and procedures related to watermarking and content provenance, it should embrace tools already used by industry and familiar to developers and deployers, which will best balance safety and security concerns with openness and innovation.

Specifically, regarding the economic implications, generative AI has the potential to produce immense benefits for individuals, businesses, and entire industries. Estimates of the economic impact of the technology are in the trillions of dollars, with some reports projecting a 7 percent increase in global gross domestic product and productivity growth by 1.5 percent over a 10-year period<sup>[4]</sup>. Ensuring that model developers have clear

standards regarding available training data, as well as clarity around the processes for risk management or auditability, will impact investment and use of these systems. Research literature illustrates the positive impact open standards can have on technological innovation, adoption, and trust among market participants.[5] For the economic benefits of the technology to come to fruition, businesses and individuals will want to be confident the models they rely on are safe, secure, and trustworthy. Technologies such as watermarking or tools to establish content provenance should be considered by NIST as one way to authenticate the origin and evolution of generative AI content. Establishing guidance that balances openness and innovation with security and trust should be central to NIST's work in this area to maximize the economic benefits.

At the same time, there are also valid concerns that advancements in AI technology absent clear guidance could result in significant economic disruption and potentially negative externalities. There is active litigation over the unauthorized use of works of books, music, movies, and other media to train Large Language Models (LLM) and the outputs of such models as noted above.[6] Some creatives say these technologies are violating copyrights by using their work to train, depriving them of revenue today, and producing outputs that substitute for their work, potentially replacing them tomorrow.[7] Model developers, notably OpenAI, have struck deals with media companies to license content for training.[8] While certain parties argue these types of agreements should become standard, such arrangements would preference large incumbents or financially connected developers over new entrants or open-source developers and could create a competitive moat for incumbents.[9] As NIST crafts guidance around acceptable use and auditability, it should avoid guidance that would limit available training data or require certain permissions that could impact competition and innovation. NIST should also consider how programs such as the National AI Research Resource and similar programs can help address some of these challenges.[10]

#### Crafting Standards to Promote Economic Benefits Generative AI

Two potential solutions to designate ownership or original creation currently being used and developed are “watermarks,” which help identify whether an image was created or modified by an AI-agent, and “content provenance,” which labels data and creates a traceable chain to understand where, when, and how a piece of content is created and changes overtime. The Coalition for Content Provenance and Authenticity (C2PA), and the World Wide Web Consortium (W3C) verifiable credential data model both provide standards and recommendations for the technical design and implementation of content provenance and watermarking for AI-generated content.[11] Google DeepMind's SynthID similarly embeds a digital watermark into AI-generated content as well as scanning of an image or audio clip for a watermark to identify if the content was synthetically generated or altered.[12] Because private firms are already deploying these tools, NIST should evaluate their real-time impact and potential for adaptation or inclusion in its guidance.

#### Crafting Standards to Promote Secure Generative AI

Further, NIST should also explore standards that promote the security of AI-powered technologies, especially in industries where technical vulnerabilities could expose personal data, compromise critical infrastructure, or contribute to the dissemination of harmful or intentionally misleading content.

NIST should focus on standards to 1) mitigate harms created by models and 2) address vulnerabilities within the model. Standards to address harms from the models should focus on scenarios such as using models to craft and scale new forms of cyberattacks or strip or manipulate provenance information from previously edited or authentic content. The latter would consider how developers and deployers can secure their systems from adversarial attacks, either to “jailbreak” the model, or undermine or compromise the software itself. NIST's Cybersecurity Framework should complement any guidance through this proceeding, especially for issues

related to network security, data management and storage, and threat evaluation.[13] Further, NIST can look to firms developing and offering foundation models such as OpenAI, Anthropic, Meta, and Google, all of which have published information on the types of red teaming used prior to their public release.[14]

Finally, in crafting guidance, NIST should implement third-party organizations' standards and tools for content provenance and watermarking as noted above. Using private standards and technologies to frame NIST's guidance could help familiarize the new standard to entities building and deploying AI-powered technologies, specifically in the context of generative AI. And because of these organizations' multi-stakeholder nature, guidance relying on these standards can help build consensus among nation-states as well as firms, increasing cross-border cooperation and trade for AI/Machine Learning (ML)-powered technologies and use-cases.

## **Red-Teaming and Information Sharing**

The RFI asks about establishing guidelines for appropriate processes and procedures to enable AI model developers to conduct AI red-teaming to enable the deployment of safe, secure, and trustworthy systems.[15] NIST guidance regarding red-teaming should incorporate existing industry and multi-stakeholder organization procedures and protocols to promote diffusion and improvements. Generative AI red-teaming focuses primarily on "prompt hacking." More akin to boundary or stress testing, prompt hacking red teamers try to overcome the guardrails input by developers into a generative AI model or LLM chatbot. This process prevents misuse or harm of the model and can limit nefarious uses by bad actors. Because the concerns regarding model misuse are varied, NIST should design flexible standards that allow for a wide range of red-teaming strategies to account for changing model architecture and corresponding security risks. Specifically, considering how different models are built and fine-tuned will have an impact on which red-teaming strategies will be most effective and any guidance from NIST should account for such variation.

While there has been a focus on red-teaming in AI to prevent harmful or false outputs, this should not overshadow attempts to address traditional cybersecurity concerns related to AI models.[16] For some firms, potential harm stemming from outputs and bypassing guardrails has been more salient than traditional security concerns and more attention is paid when discussing their red-teaming activities. The Request for Information, however, highlights the importance of red-teaming to address security concerns such as identifying threat models, specific security risks for models, and the economic feasibility for firms to conduct red-teaming. Based on the different strategies used and the different vulnerabilities being exploited, it may be more appropriate to bucket generative AI monitoring into two camps. The first is AI red-teaming, which deals with traditional intellectual property (IP), privacy, and security of AI systems. The second could be called AI testing, which focuses on safety and output of the AI tools functioning properly. This would allow for red-teaming guidance to be targeted at specific types of harm and responsive to continued evolution of model capabilities and vulnerabilities.[17] In practice, this reframing of monitoring protocols for AI models will aim to ensure that the enhanced security protocols that have become commonplace in traditional software applications are able to be translated into the fast-moving generative AI space.

The NIST framework should also encourage collaboration in red-teaming and information sharing to crowdsource security improvements among AI models. An example of a framework that should inform NIST's guidance is the Financial Services Information Sharing and Analysis Center (FS-ISAC). The FS-ISAC organization has developed standards to prioritize privacy and security for firms within the global financial system. Further, the lack of knowledge around the intricacies within more sophisticated generative AI tools and the "black box" phenomenon can hamper security improvements.[18] Outside of traditional methods to input security compliance into software, unsupervised AI tools (as opposed to supervised) with consistent human intervention present an interesting possible mechanism to aid in security protocols, as these AI tools have the

best current potential to be more resilient to new attacks on systems.<sup>[19]</sup> Examining the performance of these models through red-teaming events, such as the DefCon hacker convention, could be beneficial for identifying existing capabilities and weaknesses. NIST should build on existing collaborations such as the C2PA standards or the Partnership on AI to reflect current efforts to mitigate harm related to the development and deployment of AI models.

## **Reducing Risks of Synthetic Content**

The White House's AI executive order (EO) lays out a non-exhaustive list of ideas to address content liability risks associated with synthetically generated content, primarily regarding the generation and hosting of unlawful content and content that violates the copyright of another. With the number of organizations making these tools available growing exponentially, implementing guidelines to protect product owners of AI tools, end-users, and owners/producers of training data is paramount. NIST, however, should be careful not to impede current developments in courts and Congress.

Regarding unlawful content, there are differing views on who should ultimately bear liability for synthetic content creation and whether Section 230 largely immunizes online platforms for hosting such content. Based on recent decisions such as *Gonzalez v. Google* and scholarly pondering on the subject, it is likely that liability for unlawful content that is produced or manipulated should be bestowed upon the person prompting the AI model.<sup>[20]</sup> If an end user designs an LLM to produce or manipulate content that could hold someone liable (e.g., a terrorist group uses ChatGPT to recruit members to commit a violent act), there is 1) a long list of roadblocks that would make liability against an AI model creator/host difficult and 2) a long list of U.S. precedent regarding user-created content that aligns with this framework.<sup>[21]</sup> To the extent that policymakers find existing legal authority lacking to tackle unlawful content, Congress can introduce legislation to shift liability as appropriate. Because litigation and legislative debates are ongoing, NIST should refrain from making major recommendations related to liability and focus any guidance on standards related to transparency or identification of synthetic content generation, as discussed above.

Section 230 notably contains exceptions for intellectual property claims, and generative AI raises novel copyright infringement challenges relating to the training data on which AI-powered tools are built. OpenAI has addressed this issue regarding its leading platform, ChatGPT, by offering paid users funds when faced with copyright lawsuits against content created by the AI.<sup>[22]</sup> But OpenAI's solution is a short term one that puts a Band-Aid on a much larger problem. So long as models run on datasets that include copyrighted and protected data, problems in copyright liability will continue to fester. As noted previously, there is active litigation focused on whether certain training and outputs of AI models violated copyright of authors, musicians, and artists. The U.S. copyright and patent office held a proceeding to consider updating its rules to adapt to the emerging challenges presented by generative AI.<sup>[23]</sup> Similarly, Congress has also held multiple hearings and members have introduced bills addressing some of the issues related to generative AI and copyright. Until that occurs, encouraging and recommending cross-industry dialogue between model builders and owners of copyrighted material in data sets, and recommending thresholds of copyrighted data in models, could be of use going forward to build trust in models and spur use without restraint in a broader population.

In initiating guidance to mitigate potential harms of synthetic content, NIST should allow courts and legislators to lead, but continue emphasizing common-sense standards and recommendations that aid model owners and crafters of generative AI outputs. Guidance should focus on further industry adoption and acceptance of standards, as this could help spur congressional approval on federal regulatory agencies to put into place current and future NIST guidance.

## Responsible Global Technical Standards

### Crafting Standards to Promote Innovation and Trust Globally

The third and final component of the RFI asks how NIST can develop guidance to promote AI-related standards that build some consensus with other nation-states and international bodies. The AI EO requires the secretaries of the Department of Commerce and State to establish a plan for global engagement on promoting and developing AI consensus standards that align with ideas presented in the NIST AI RMf and the U.S. Government National Standards Strategy for Critical and Emerging Technology (NSS-CET).<sup>[24]</sup> In addition to the recommendations from this report, NIST should use the ideas that influenced the RMF and NSS-CET as foundations for certain standards and definitions.

With regard to questions of nomenclature and other definitions, as well as best practices for data and model training, the NIST AI RMF should be a foundational document. The NIST RMF's focus on managing risk and delineating the risks commensurate with the use of certain AI-powered technologies makes it an ideal document for guiding international standards on related topics. The RMF is designed to be iterative, which allows it to respond to evolutions in the development and deployment of AI/ML technologies. The RMF itself draws from previous work done by the Organization for Economic Cooperation and Development (OECD), whose work has been influential in the drafting of strategies and legislation in several member countries, including the United Kingdom, the European Union (EU), and Japan.<sup>[25]</sup> The document's focus on mitigating risks based on the design and deployment of technology, its iterative character, and common foundation with other international partners make the RMF an important component of any attempt to develop consensus standards.

NIST should also use the NSS-CET as a reference point for considering the trade-offs of standards and guidance related to data access and management, trustworthiness and verification, and the implications of standards for competition and international trade. Specifically, the document calls for collaboration on standards, noting the important role the private sector and non-governmental organizations can play. As mentioned above, for issues related to verification and content provenance, several organizations composed of firms and civil society are developing technologies and protocols to address such issues. These solutions should inform NIST's approach to such areas. Additionally, the NSS-CET notes the important role standards can play in driving investment and research and development. This is an area where harmonization with allied nations can promote foreign investment and collaboration with firms, academia, and civil society.

### Weighing the Tradeoffs Related to Global Consensus Standards

International collaboration and harmonization can have many benefits, but NIST and other relevant stakeholders should be cognizant of the tradeoffs created by regulatory first movers, particularly the EU. Similar to the script of the General Data Protection Regulation (GDPR), the EU is attempting to pass a sweeping law regulating the development and deployment of AI, which could allow it to set a colloquial "floor" for AI regulation globally.<sup>[26]</sup> As NIST considers the implications of incongruence with the EU, it should pay particular attention to how the EU's proposal could impact barriers to entry, new business formation, and compliance costs for small and medium-sized firms. Research has shown the harmful effects GDPR has had on new business formation and competition within digital markets in the EU.<sup>[27]</sup> As NIST evaluates other nations' approaches to standards and guidance, it should be particularly wary of the EU's approach.

In addition to the EU, the People's Republic of China (PRC) has enacted and begun implementing regulations governing the types of data and systems that can be built and adhere to the nation's ideological censorship



restrictions.[28] The PRC has made a point to contribute to and influence other standards bodies, particularly in the field of telecommunications. The United States and other nations, such as the UK, Japan, and Israel, have all put out documents and plans for an approach to standards and governance that emphasize a flexible, light-touch approach. Many of these ideas can be found within work done by the OECD, which can be a helpful starting point for collaboration abroad. International collaboration can be beneficial, but NIST should be cognizant of how actors may use such forums to advance agendas antithetical to the United States' goals in AI/ML development and deployment.

## Conclusion

If approached wisely, NIST can lay the foundation for standards that could help promote innovation and competition within the market for AI-powered technologies. By providing guidance for developers and deployers on issues related to model development, red-teaming and stress testing, appropriate data practices, and many others, NIST can encourage firms and individuals to build models that are safe, secure, and trustworthy. Further, as guidance is developed, looking to the private sector, multi-stakeholder organizations, and other nations can provide useful references in areas such as watermarking and content provenance. To achieve the goal outlined by AI EO and the RFI, NIST should consider how previous documents and standards it and other bodies have promulgated can contribute to the present and future iterations of guidance.

[1] Artificial Intelligence Risk Management Framework, NIST AI 100-1, (2023); The NIST Cybersecurity Framework 2.0, NIST, (2023).

[2] Jeffrey Westling, Are Deep Fakes a Shallow Concern? A Critical Analysis of the Likely Societal Reaction to Deep Fakes, 2-12 (2019).

[3] Nabila Ahmed, Adam Haigh, Ainsley Thomson, and Ellie Harmsworth, *Deepfake Imposter Scams Are Driving a New Wave of Fraud*, (2023), <https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime?sref=S0MJebS0>

[4] Michael Chui, et. Al, *The Economic Potential of Generative AI: The Next Productivity Frontier*, McKinsey Digital, (2023), <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction>; *Generative AI Could Raise Global GDP by 7%*, Goldman Sachs, (2023), <https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html>

[5] Jorge Padilla, John Davies, and Aleksandra Boutin, *Economic Impact of Technology Standards*, 11-22, 26-30, 32-40, (2017), [https://www.compasslexecon.com/wp-content/uploads/2018/04/CL\\_Economic\\_Impact\\_of\\_Technology\\_Standards\\_Report\\_FINAL.pdf](https://www.compasslexecon.com/wp-content/uploads/2018/04/CL_Economic_Impact_of_Technology_Standards_Report_FINAL.pdf); Knut Blind, Florian Ramel, and Charlotte Rochell, *The Influence of Standards and Patents on Long-Term Economic Growth*, 981-84, 987-995, (2021), <https://link.springer.com/article/10.1007/s10961-021-09864-3>; Sujai Shivakumar, *Securing Global Standards for Innovation and Growth*, CSIS, (2022), <https://www.csis.org/analysis/securing-global-standards-innovation-and-growth>; Ira Kallish, Michael Wolf, Jonathan Holdowsky, *The Link Between Trust and Economic Prosperity*, Deloitte, (2021), <https://www2.deloitte.com/us/en/insights/economy/connecting-trust-and-economic-growth.html>.

[6] Joshua Levine, John Belton, *Primer: Training AI Models with Copyrighted Work*, American Action Forum,

(2023), <https://www.americanactionforum.org/insight/primer-training-ai-models-with-copyrighted-work/>.

[7] Creative Economy and Generative AI, FTC, (October 4, 2023), <https://www.ftc.gov/news-events/events/2023/10/creative-economy-generative-ai>; Artificial Intelligence and Intellectual Property – Part II: Copyright, U.S. Senate Committee on the Judiciary, Subcommittee on Intellectual Property, (July 12, 2023), [https://www.judiciary.senate.gov/artificial-intelligence-and-intellectual-property\\_part-ii-copyright](https://www.judiciary.senate.gov/artificial-intelligence-and-intellectual-property_part-ii-copyright).

[8] Partnership with Axel Springer to Deepen Beneficial Use of AI in Journalism, OpenAI, (December 13, 2023), <https://openai.com/blog/axel-springer-partnership>.

[9] Oversight of A.I.: The Future of Journalism, U.S. Senate Committee on the Judiciary, Subcommittee on Privacy, Technology, and the Law, (January 10, 2024), <https://www.judiciary.senate.gov/committee-activity/hearings/oversight-of-ai-the-future-of-journalism>; *Big Ideas* 2023, 05 – 15, 20-29, 30-39, ARK Invest, (January 31, 2023), [https://research.ark-invest.com/hubfs/1\\_Download\\_Files\\_ARK-Invest/Big\\_Ideas/ARK%20Invest\\_013123\\_Presentation\\_Big%20Ideas%202023\\_Final.pdf](https://research.ark-invest.com/hubfs/1_Download_Files_ARK-Invest/Big_Ideas/ARK%20Invest_013123_Presentation_Big%20Ideas%202023_Final.pdf).

[10] Democratizing the future of AI R&D: NSF to launch National AI Research Resource Pilot, U.S. National Science Foundation, (January 24, 2023), <https://new.nsf.gov/news/democratizing-future-ai-rd-nsf-launch-national-ai>.

[11] Coalition for Content Provenance and Authenticity, <https://c2pa.org/>; World Wide Web Consortium (W3C), <https://www.w3.org/standards/>.

[12] SynthID, <https://deepmind.google/technologies/synthid/>

[13] The NIST Cybersecurity Framework 2.0, NIST, (2023).

[14] GPT-4 System Card, OpenAI, (March 23, 2023), <https://cdn.openai.com/papers/gpt-4-system-card.pdf>; *Why Red Teams Play a Central Role in Helping Organizations Secure AI Systems*, Google, (July, 2023), [https://services.google.com/fh/files/blogs/google\\_ai\\_red\\_team\\_digital\\_final.pdf](https://services.google.com/fh/files/blogs/google_ai_red_team_digital_final.pdf); *Red Teaming Language Models to Reduce Harms: Methods, Scaling Behaviors, and Lessons Learned*, Anthropic, (August, 22, 2022), <https://www.anthropic.com/news/red-teaming-language-models-to-reduce-harms-methods-scaling-behaviors-and-lessons-learned>; *Llama 2: Open Foundation and Fine-Tuned Chat Models*, Meta, (July 18, 2023), <https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/>.

[15] Jessica Ji, *What Does AI Red-Teaming Actually Mean?* Center for Security and Emerging Technology, (October 24, 2023), <https://cset.georgetown.edu/article/what-does-ai-red-teaming-actually-mean/>.

[16] Alan Mislove, *Red-Teaming Large Language Models to Identify Novel AI Risks*, Office of Science and Technology Policy, (August 29, 2023), <https://www.whitehouse.gov/ostp/news-updates/2023/08/29/red-teaming-large-language-models-to-identify-novel-ai-risks/>.

[17] Jessica Ji, *What Does AI Red-Teaming Actually Mean?* Center for Security and Emerging Technology, (October 24, 2023), <https://cset.georgetown.edu/article/what-does-ai-red-teaming-actually-mean/>.

[18] Sunil Ramlochan, *The Black Box Problem: Opaque Inner Workings of Large Language Models*, Prompt Engineering Institute, (October 23, 2023), <https://promptengineering.org/the-black-box-problem-opaque-inner-workings-of-large-language-models/>

- [19] Seth Lloyd, Masoud Mohseni, Patrick Reber, *Quantum algorithms for supervised and unsupervised machine learning*, (November 4, 2013), <https://arxiv.org/abs/1307.0411>.
- [20] Eugene Volokh, Mark A. Lemley, and Peter Henderson, *Freedom of Speech and AI Output*, 651-54, *Journal of Free Speech Law*, (2023), <https://www.journaloffreespeechlaw.org/volokhlemleyhenderson.pdf>
- [21] Ibid; Peter Henderson, *Who is Liable When Generative AI Says Something Harmful?* Stanford University Human-Centered Artificial Intelligence, (October 11, 2023), <https://hai.stanford.edu/news/who-liable-when-generative-ai-says-something-harmful#:~:text=Those%20seeking%20to%20impose%20liability,in%20all%20of%20these%20cases>
- [22] Blake Montgomery, *OpenAI Offers to Pay for ChatGPT Customers' Copyright Lawsuits*, The Guardian, (November 6, 2023), <https://www.theguardian.com/technology/2023/nov/06/openai-chatgpt-customers-copyright-lawsuits>
- [23] United States Copyright Office, Copyright Registration Guidance: Works Containing material Generated by Artificial Intelligence, 16190 Federal Register, Vol. 88, No. 51, 37 CFR Part 202, <https://www.copyright.gov/ai/>
- [24] Joseph R. Biden Jr., Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, (October 30, 2023), <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>; Artificial Intelligence Risk Management Framework, NIST AI 100-1, (2023); United States Government National Standard Strategy for Critical and Emerging Technology, (May 2023), <https://www.whitehouse.gov/wp-content/uploads/2023/05/US-Gov-National-Standards-Strategy-2023.pdf>.
- [25] OECD, *Recommendation of the Council on Artificial Intelligence*, (July 7, 2023), <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>; AI Regulation: A Pro-Innovation Approach, Department for Science, Innovation and Technology and Office for Artificial Intelligence, United Kingdom, (March 29, 2023), <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>; Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>; Governance Guidelines for Implementation of AI Principles, Ver. 1.1, Expert Group on How AI Principles Should be Implemented, (January 28, 2022), [https://www.meti.go.jp/shingikai/mono\\_info\\_service/ai\\_shakai\\_jisso/pdf/20220128\\_2.pdf](https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20220128_2.pdf); Social Principles of Human-Centric AI, <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>.
- [26] Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>
- [27] Rebecca Janßen, Reinhold Kesler, Michael E. Kummer & Joel Waldfogel, *GDPR and the Lost Generation of Innovative Apps*, NBER, (May 2022), <https://www.nber.org/papers/w30028>; Garrett A. Johnson, *Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond*, NBER, (June 16, 2023), <https://www.nber.org/system/files/chapters/c14784/c14784.pdf>; Carl Benedikt Frey, Giorgio Presidente, *The GDPR effect: How data privacy regulation shaped firm performance globally*



, VoxEU, (March 10, 2022), <https://cepr.org/voxeu/columns/gdpr-effect-how-data-privacy-regulation-shaped-firm-performance-globally>

[28] Cybersecurity Administration of China, Provisions on the Administration of Deep Synthesis Internet Information Services & Administrative Measures for Generative Artificial Intelligence, (December 11, 2022), [http://www.cac.gov.cn/2022-12/11/c\\_1672221949354811.htm](http://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm); Cybersecurity Administration of China, Notice of the Cyberspace Administration of China on the Public Solicitation of Opinions on the “Measures for the Administration of Generative Artificial Intelligence Services (Draft for Comment),” (April 11, 2023), [http://www.cac.gov.cn/2023-04/11/c\\_1682854275475410.htm](http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm).