



# The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction

Emmie Hine<sup>1</sup> · Luciano Floridi<sup>1,2</sup>

Received: 17 November 2022 / Accepted: 14 January 2023 / Published online: 29 January 2023  
© The Author(s) 2023

## Abstract

The US is promoting a new vision of a “Good AI Society” through its recent AI Bill of Rights. This offers a promising vision of community-oriented equity unique amongst peer countries. However, it leaves the door open for potential rights violations. Furthermore, it may have some federal impact, but it is non-binding, and without concrete legislation, the private sector is likely to ignore it.

**Keywords** AI governance · AI policy · US AI policy · AI Bill of Rights · Comparative policy · Blueprint for an AI Bill of Rights

In October of 2022, the White House Office of Science and Technology Policy released the “Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People” (the “Blueprint”) after a yearlong public engagement process. When the process began in 2021, the stated goal was for “a bill of rights for an AI-powered world” (Lander, 2022), but the finished product, described as a “white paper” and a “framework” (White House Office of Science and Technology Policy, 2022a), is a much less ambitious, non-binding guide to the office’s vision for AI development. Although the principles it lays out are not yet enforceable, it is the clearest look yet at the Biden administration’s definition of a “Good AI Society” (Cath et al., 2018). In this Comment, we analyze that vision and compare it to the two most prominent players on the global AI stage, the European Union (EU) and China. We argue that it is a unique and promising, if flawed, advance in US AI governance, but requires concrete legislation to be effective in the private sector.

---

✉ Emmie Hine  
emma.hine@studio.unibo.it

<sup>1</sup> Department of Legal Studies, University of Bologna, Via Zamboni, 27, 40126 Bologna, Italy

<sup>2</sup> Oxford Internet Institute, University of Oxford, 1 St. Giles’, OX1 3JS Oxford, UK

## 1 An Unenforceable but Promising Start

As a non-binding white paper, the Blueprint outlines principles and ideas for how to put them into practice, rather than concrete requirements and actions. It shares with the EU's Artificial Intelligence Act (AIA) the same product-oriented approach to AI as a system, not as a source of outputs, which then should be regulated. However, instead of attempting to define AI, as the EU is struggling to do in the AIA, the Blueprint applies to “automated systems” that could “meaningfully impact” the American public’s civil rights, civil liberties, and privacy; access to equal opportunities; or “access to critical resources or services” (White House Office of Science and Technology Policy, 2022a). This is a step in the right direction, that is, towards regulating the uses of AI, not AI itself. However, it is insufficiently bold in its approach to implementation, and the appendix definition of an “automated system” is too inclusive: “any system, software, or process that uses computation as whole or part of a system to determine outcomes, make or aid decisions, inform policy implementation, collect data or observations, or otherwise interact with individuals and/or communities” (White House Office of Science and Technology Policy, 2022a). This could describe almost any system with which a user interacts, as computation is inseparable from software and hardware. However, the requirement that the impact is “meaningful” would limit the automated systems under consideration to AI systems in critical applications, provided the “meaningful” specification itself is meaningful. This resembles some of the modifications the Czech Presidency recently suggested to the AIA (Bertuzzi, 2022a).

The Blueprint outlines five principles:

- Safe and effective systems.
- Protection from algorithmic discrimination and inequitable systems.
- Protection from abusive data practices and agency about how personal data is used.
- Knowledge of when an automated system is used and understanding of its impacts.
- The ability to opt out of automated systems for a human alternative.

For each principle, the Blueprint lays out its importance and what should be expected of automated systems to conform to the principle, and outlines how it can be put into practice through broad descriptions of examples. It contains no enforcement mechanisms and thus has been criticized as “toothless” (Johnson, 2022). Still, as a strategic document, it may have a federal impact even without legislative efforts to codify its ambitions.

## 2 An Evolution of the US Good AI Society

Each administration since Obama has suggested a new idea of a Good AI Society. Roberts et al. (2022) evaluate AI governance approaches by looking at *how* governments enact their goals and *whom* policies are intended to benefit. This framework

helps assess the Blueprint and compare its vision to others. In terms of whom is to benefit, the vision laid out in the Blueprint focuses on empowering communities through protecting individual rights. Noting that existing frameworks offer inadequate protections when impacts manifest at the community level, the Blueprint provides a broad definition of community that includes both neighborhoods, on- and offline social connections, formal organizational ties, and affinity- or identity-based connections (White House Office of Science and Technology Policy, 2022a). It states that harms perpetuated by automated systems must be assessed and redressed at both the individual and community level. This represents a new avenue in the US AI vision. Furthermore, in writing in the second person, the Blueprint deliberately addresses readers as members of the American public rather than users or consumers. In its focus on equity, countering discrimination, and listening to public input, the Blueprint incorporates aspects of the Obama era of AI regulation (Hine & Floridi, 2022), but its explicit focus on community-level equity is new and unique internationally. Unfortunately, like the AIA, there is no mention of the environment, as if AI did not have an impact on it, and such impact had no significance for society (Cowls et al., 2021). A broad interpretation of community impact could include climate change, but the latter does not seem to be a priority in the Blueprint.

The Obama administration issued several AI-related documents in its final year, outlining a diversity-focused approach aiming to develop AI “*by and for*” diverse populations (Executive Office of the President, 2016). The Blueprint focuses less on increasing the diversity of the development pipeline and more on ensuring that different communities’ voices are heard throughout the development process, mirroring the year of public input that shaped the Blueprint. This outlines a focus on equity not predicated on the slow reshaping of the software industry. The Obama era also began emphasizing a faith in American innovation that continued into the Trump administration, which centered its nebulous Good AI Society vision on promoting ill-defined “American values” domestically and abroad for the benefit of the US and its allies, but not its main competitor, China (Hine & Floridi, 2022). At the beginning of Biden’s term, his administration’s approach was more in line with the ethics-centered rhetoric of the Obama era, but retained aspects of the Trump era’s emphasis on trustworthy AI – also referenced in the Blueprint, though perhaps rendered irrelevant by its new principles – and competition with China (Hine & Floridi, 2022). The Blueprint, however, is focused on the American people, rather than US enemies and allies.

Regarding how the vision will be enacted, federal departments announced action across labor, consumer protection, education, healthcare, and housing (White House Office of Science and Technology Policy, 2022b). Some of these, such as a collaboration between the Department of Labor and the Equal Employment Opportunity Commission (EEOC) on reimagining hiring and recruiting practices, including the use of automated systems, or a crackdown on algorithmic discrimination in the financial system, could have tangible impact. However, as of early 2023, reports by the Department of Health and Human Services on health equity and the Department of Energy on responsible AI due in 2022 have yet to materialize, raising questions about follow-through. Furthermore, private sector action is unlikely without further legislative action. The Blueprint focuses mainly on rights and less on duties and on promoting new AI development, but has an implicit faith in “the power of Ameri-

can innovation” woven throughout (White House Office of Science and Technology Policy, 2022a). Unlike in the Trump administration, this faith is more cautious. The Blueprint notes that automated systems can drive innovation, but also have the potential to embed discrimination (White House Office of Science and Technology Policy, 2022a). Thus, the Blueprint can be seen as a set of guardrails preventing the unfettered power of American innovation from causing negative externalities rather than a motor spurring new developments.

The Blueprint itself does not concern itself with international cooperation or competition, although soon after its release, new sanctions targeting the Chinese semiconductor industry were announced (Brands, 2022), which continues the escalation conducted over the Trump and the start of the Biden administrations (Hine & Floridi, 2022). It is internally focused and more concerned with encouraging the development of domestic, industry-specific standards rather than influencing international standards. The most significant external effort – under the header “Leading by example and advancing democratic values” in the fact sheet accompanying its release – is the launch of an AI Action Plan for the United States Agency for International Development, a development assistance organization. Even the document acknowledges that this is to “sustain American global leadership” rather than advance it (White House Office of Science and Technology Policy, 2022b). This domestic focus is perhaps appropriate for a document that brands itself as a “Bill of Rights,” as the original Bill of Rights was written to guarantee the rights of the American people.<sup>1</sup> Regardless, this turning inwards is a notable development.

### 3 Comparing to AI Peers: Promise and Concerns

A fruitful way to understand the nature and implications of the community-oriented Blueprint is to compare it to the Good AI Society visions of the two other major AI powerhouses, the EU and China. The EU’s Good AI Society takes an “explicitly individual-focused,” human-centric approach, which involves promoting development in the public and private sectors and coordinating across Member States (Roberts et al., 2022). China’s approach is highly innovation-focused, with initiatives to promote specific developments at all levels of government and industry, and societally oriented, prioritizing social stability over the rights of the individual (Hine & Floridi, 2022; Roberts et al., 2022).

Since the Blueprint is not a binding document, it will not independently enact the administration’s goals. Throughout, it acknowledges the need for more concrete legislation and hints at how its vision might be brought to life. It calls for a comprehensive data privacy law rather than the “patchwork” that exists today, acknowledging data privacy as “fundamental” to enabling the other principles (White House Office of Science and Technology Policy, 2022a). Many of the ideas it proposes – like

---

<sup>1</sup> With the exception of the “right to not be discriminatorily denied the vote” and the right to run for federal office, the Constitution and Bill of Rights were written to apply to “the people” regardless of citizenship (Cole, 2010). Even in its extended title, the Blueprint echoes this language, showing its intent to apply citizens and noncitizens alike.

meaningful consent to data collection and the right to correct erroneous personal data – are similar to clauses in the EU’s 2018 General Data Protection Regulation (GDPR), or two of its relatives, China’s 2021 Personal Information Protection Law (PIPL) (Swabey, 2021) and the California Consumer Privacy Act (CCPA) (Keane, 2021). One stated aim of the Blueprint is that the public should be protected from “unchecked surveillance,” but it has been criticized for normalizing surveillance in a manner more akin to China than the EU (Fox Cahn, 2022). The EU’s AIA may ban both real-time and ex-post biometric identification via surveillance devices, except to identify victims or perpetrators of crime or to prevent an imminent threat (Bertuzzi, 2022b). China’s PIPL, on the other hand, permits mass surveillance in the name of national security (Roberts et al., 2022). The Blueprint states that “Surveillance should be avoided unless it is strictly necessary to achieve a legitimate purpose and it is proportionate to the need” and that surveillance must not limit civil liberties or civil rights (White House Office of Science and Technology Policy, 2022a). However, the broad language leaves the door open to surveillance, including real-time surveillance in public areas, with a potential negative impact on segments of the population already subject to heightened nontechnical surveillance. Furthermore, while the AIA would ban AI applications deemed unacceptably risky and put increased requirements on high-risk systems, the Blueprint does not forbid any applications (Eliot, 2022). It only states that systems that violate its principles should not be developed. Thus, although this new Good AI Society is meant to support individuals and communities, the latitude in its principles may mean that mass surveillance and unethical systems could increase.

One promising effort is in algorithmic explanation. The Blueprint states that users should be provided “generally accessible plain language documentation” outlining the system, the role of automation, and an explanation of how an outcome was reached. The GDPR ostensibly enshrines a right to “meaningful information” about the logic behind a decision, but the ambiguity of the provisions raises substantial questions about the scope of the “right to explanation” and its feasibility in practice (Kaminski, 2019; Wachter et al., 2017). China’s “Provisions on the Management of Algorithmic Recommendations in Internet Information Services” requires only an explanation when “rights and interests” are majorly impacted (China Law Translate, 2022). The Blueprint’s general specifications could balance technical feasibility with information usability and entitle individuals to truly meaningful explanations of algorithmic decisions.

However, some of its more specific provisions could be technically infeasible. On top of its general explanation requirements, the Blueprint specifies that only “fully transparent,” explainable models should be used in areas where “extensive oversight is expected,” such as criminal justice. It is unclear what “fully transparent” means, and many algorithms that rely on machine learning are inherently inscrutable. China’s growing use of AI in courts has raised concerns around the explainability of decisions and appropriateness of training data (Stern et al., 2021). The AIA considers AI in the criminal justice system to be high-risk and thus places it under the same data quality requirements as other high-risk applications. However, it does not impose additional explainability requirements beyond what the GDPR vaguely contains – users merely have to be able to “interpret the system output and use it appropriately” (Permanent

Representatives Committee, 2022). The three different standards could lead to substantially different uses of AI in criminal justice that could differentially threaten individual rights. However, the US's approach risks precluding any possible benefits.

A major component of national AI governance strategies is promoting innovative developments in AI. The Blueprint does not outline explicit measures to promote innovation. It does not take a "wish list" approach to developments like China Hine & Floridi, 2022. Nor does it encourage the formation of regulatory sandboxes or increase SME support like the AIA. Indeed, the accompanying Fact Sheet's entrepreneur support measures are a forthcoming risk management framework and a *continuation* of existing National Science Foundation funding, leaving the innovation landscape essentially unchanged. It is also unclear whether any auditing-like approach, as present in the AIA, could be created based on the Blueprint.

#### 4 Unequal Impacts on the Public and Private Sectors

The Blueprint is already being promoted across the federal government. Its wide-ranging efforts could have significant impacts on aspects of American life. Some industry actors could adopt measures from the Blueprint without further incentive, and the frameworks being developed by federal agencies could trickle down to industry. However, the Blueprint is unlikely to impact the private sector unless, as it suggests, additional legislation is advanced (Johnson, 2022). As a "Bill of Rights," it could be much more ambitious about spurring action. A bill of rights is meant to guarantee a set of rights to specific people (Merriam-Webster, n.d.); it must, by definition, be enforceable. The original Bill of Rights did not begin as a "blueprint," but as a way to define and guarantee Constitutional freedoms at the dawn of a new era in America. For the Blueprint to do the same in the era of AI, it must have teeth. More of the administration's vision may become a reality if the Blueprint spurs action on the stalled Algorithmic Accountability Act of 2022 or the American Data Privacy and Protection Act. Another avenue for impact would be if it inspires state legislation on the level of the CCPA or state-level biometric privacy laws, such as the Illinois Biometric Information Privacy Act, under which Facebook paid \$550 million to settle a class-action lawsuit regarding its use of facial recognition without notification (Pester, 2020). The Blueprint represents a promising (if significantly improvable) development of the American Good AI Society that emphasizes the protection of communities as well as individuals. However, without concrete legislation to back it, it risks remaining a mere invitation.

**Funding** Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

#### Declarations

**Competing Interests** The authors have no competing interests to declare.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long

as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bertuzzi, L. (2022a, September 16). AI Act: Czech Presidency puts forward narrower classification of high-risk systems. *Www.Euractiv.Com*. <https://www.euractiv.com/section/digital/news/ai-act-czech-presidency-puts-forward-narrower-classification-of-high-risk-systems/>
- Bertuzzi, L. (2022b, October 6). AI Act: EU Parliament's discussions heat up over facial recognition, scope. *Www.Euractiv.Com*. <https://www.euractiv.com/section/digital/news/ai-act-eu-parliaments-discussions-heat-up-over-facial-recognition-scope/>
- Brands, H. (2022, October 9). Biden's Chip Limits on China Mark a War of High-Tech Attrition. *Bloomberg.Com*. <https://www.bloomberg.com/opinion/articles/2022-10-09/biden-s-chip-limits-on-china-mark-a-war-of-high-tech-attrition>
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the 'Good society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24, <https://doi.org/10.1007/s11948-017-9901-7>.
- China Law Translate (2022, January 4). Provisions on the Management of Algorithmic Recommendations in Internet Information Services 互联网信息服务算法推荐管理规定. *China Law Translate*. <https://www.chinalawtranslate.com/algorithms/>
- Cole, D. (2010). *Are foreign nationals entitled to the same constitutional rights as Citizens?*. Georgetown Law Faculty Publications.
- Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2021). *The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations*. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-021-01294-x>.
- Eliot, L. (2022, October 5). US Releases an AI Bill of Rights That, Though Encouraging, Won't Yet Move the Needle—JURIST - Commentary—Legal News & Commentary. <https://www.jurist.org/commentary/2022/10/u-s-releases-an-ai-bill-of-rights-that-though-encouraging-wont-yet-move-the-needle/>
- Executive Office of the President (2016). *Preparing for the Future of Artificial Intelligence*. [https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf)
- Fox Cahn, A. (2022, October 4). I respect the folks at @WhiteHouse OSTP who worked on this, but I couldn't disagree more. This is a blueprint for normalizing and accelerating AI surveillance, not combatting it. This ignores the way AI fuels discrimination and oppression in the real world. <https://whitehouse.gov/ostp/ai-bill-of-rightshttps://t.co/JvYasGjJOK> [Tweet]. *Twitter*. <https://twitter.com/FoxCahn/status/1577297762818969601>
- Hine, E., & Floridi, L. (2022). Artificial intelligence with American values and Chinese characteristics: A comparative analysis of American and Chinese governmental AI policies. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-022-01499-8>
- Johnson, K. Biden's AI Bill of Rights Is Toothless Against Big Tech., & *Wired* (2022, October 4). <https://www.wired.com/story/bidens-ai-bill-of-rights-is-toothless-against-big-tech/>
- Kaminski, M. E. (2019). *The Right to Explanation, Explained*. <https://doi.org/10.15779/Z38TD9N83H>
- Keane, J. (2021, April 8). From California to Brazil: Europe's privacy laws have created a recipe for the world. *CNBC*. <https://www.cnbc.com/2021/04/08/from-california-to-brazil-gdpr-has-created-recipe-for-the-world.html>
- Lander, E. (2022, October 8). Americans Need a Bill of Rights for an AI-Powered World. *Wired*. <https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/>
- Merriam-Webster. (n.d.). Definition of BILL OF RIGHTS. *Merriam-Webster.Com Dictionary*. Retrieved 6 January 2023, from <https://www.merriam-webster.com/dictionary/bill+of+rights>

- Permanent Representatives Committee (2022). Proposal for a 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—General approach. <https://doi.org/10.5040/9781782258674>
- Pester, R. (2020, February 14). Patel v. Facebook: Facebook Settles Illinois Biometric Information Privacy Act (“BIPA”) Violation Suit. Harvard Journal of Law & Technology. <https://jolt.law.harvard.edu/digest/patel-v-facebook-facebook-settles-illinois-biometric-information-privacy-act-bipa-violation-suit>
- Roberts, H., Cows, J., Hine, E., Morley, J., Wang, V., Taddeo, M., & Floridi, L. (2022). Governing artificial intelligence in China and the European Union: comparing aims and promoting ethical outcomes. *The Information Society*, 0(0), 1–19. <https://doi.org/10.1080/01972243.2022.2124565>.
- Stern, R. E., Liebman, B. L., Roberts, M., & Wang, A. Z. (2021). *Automating Fairness? Artificial Intelligence in the Chinese Court* (p. 59). COLUMBIA JOURNAL OF TRANSNATIONAL LAW.
- Swabey, P. (2021, August 24). Here’s what ‘China’s GDPR’ means for international businesses. Tech Monitor. <https://techmonitor.ai/policy/heres-what-pipl-china-gdpr-means-for-international-businesses>
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76–99. <https://doi.org/10.1093/idpl/ix005>
- White House Office of Science and Technology Policy (2022a, October). Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People. The White House. <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>
- White House Office of Science and Technology Policy (2022b, October 4). FACT SHEET: Biden-Harris Administration Announces Key Actions to Advance Tech Accountability and Protect the Rights of the American Public. The White House. <https://www.whitehouse.gov/ostp/news-updates/2022/10/04/fact-sheet-biden-harris-administration-announces-key-actions-to-advance-tech-accountability-and-protect-the-rights-of-the-american-public/>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.