cogitatio

Governing AI Decision-Making: Balancing Innovation and Accountability

David Mark and **John Morison**

School of Law, Queen's University Belfast, UK

Correspondence: David Mark (dmark02@qub.ac.uk)

Abstract

This article explores the growing use of algorithmic models to make or inform decisions within the public sector. Amidst a climate of accelerating investment, expanding system applicability, and rapid technical progress, it concentrates on how key jurisdictions, most prominently the "digital empires" of the United States, European Union, and China, construct the problems associated with such algorithmic systems, and how these constructions impact governance. Drawing on an example from the legal sphere, it highlights both the potential efficiency gains and the increasing tensions concerning automation and fairness. This article then adopts aspects of Carol Bacchi's Foucauldian-inspired "What's the Problem Represented to Be?" framework to trace how divergent problem framings, ranging from the United States' emphasis on an "innovation gap," to the European Union's "trust deficit," and China's "stability risk," have produced distinct regulatory trajectories. Yet, despite these divergent framings and national strategies, this article argues that a common post-2024 trend emerges, revealing a general shift toward regulatory softening, one that privileges innovation over precautionary safeguards. This convergence raises critical questions about the future direction and resilience of "algorithmic decision-making" governance.

Keywords

accountability; AI regulation; algorithmic decision making; judicial AI; problematization; public sector innovation; techlash

1. Introduction

As governments worldwide seek to harness increasingly sophisticated algorithmic systems to complement decision-making processes, several strategic objectives are being pursued. These include ambitions for public

service transformation, the pursuit of economic growth, and efforts to secure national security and geopolitical advantage. This article investigates how key jurisdictions, specifically the US, the EU, and China, frame the risks and opportunities associated with such technologies, and how these divergent framings shape the regulation of the arena.

This article begins by explaining its choice of "algorithmic decision making" (ADM) as the central analytical terminology, before exploring the surrounding techno-economic landscape, highlighting a resurgence in algorithmic deployment, investment, and technical progress. It then examines the context in which ADM is being implemented in the public sector, using the judicial sphere to provide an illustrative example. Finally, it draws upon Carol Bacchi's "What's the Problem Represented to Be?" (WPR) framework (Bacchi, 2009) to ask how algorithmic technologies are problematized within current governance strategies and policy discourse advanced by the three key global actors, how these problem representations materialise, and what rationalities underpin them. In doing so, this article foregrounds how problem representations, ranging from an "innovation gap" in the US, a "trust deficit" in the EU, and a "stability risk" in China, operate to shape regulations. The analysis argues that these framings reflect underlying political rationalities such as market freedom, rights protection, or state control; usher in institutional logics such as precautionary legal principles or permissionless innovation; and signal techno-economic framings, such as market competitiveness or ideological stability. At the same time, somewhat paradoxically, this article identifies a recent convergence toward regulatory softening: despite divergent framings, deregulatory pressures across all three jurisdictions increasingly privilege innovation over precautionary safeguards.

The analysis draws on qualitative document analysis, centring on public policy documents, white papers, regulatory proposals, and government statements produced between 2020 and early 2025 within the three jurisdictions under study. These were selected using purposive sampling, based on their prominence in public discourse and their significance in shaping regulatory trajectories.

2. The Centrality of Decision Making

As noted, this article focuses on ADM. There are two principal reasons for this. First, operational decision-making constitutes the "most common purpose of deployed AI across government bodies" (House of Lords, 2024), making it analytically indispensable. Second, and more critically, the ADM nomenclature provides conceptual inclusivity, encompassing any system in which decision-making authority is partially or fully delegated to algorithms. Such terminology deliberately shifts attention from the technical architecture of systems to their endpoint: the juncture where data aggregation, algorithmic processing, and interpretive analysis converge to produce actionable outcomes (Yeung, 2017). The computational processes underpinning these models range from simple rule-based systems—like the British Home Office's rules-based ADM for immigration decisions (Booth, 2024a) or US federal agencies' Risk Classification Assessment systems (Department of Homeland Security, 2024)—to more complex machine learning algorithms, such as the UK Department for Work and Pensions' ML model for universal credit risk assessment (Department for Work and Pensions, 2024) and the Netherlands Tax and Customs administration's adaptive ML systems (Heikkilä, 2022). Rather than dissecting technical specificities, this analysis prioritises their unifying decision-making capacity. This approach also aligns with real-world policy and governance considerations. Many national regulatory policies tend to focus on the consequences of AI-driven decisions, rather than the particularities of the underlying technology. For example, the

classification rules in Article 6(3) and Recital 53 of Europe's seminal AI Act (Regulation (EU) 2024/1689, 2024) point to the importance of the decision-making function in assessing whether a system qualifies as high-risk. Similarly, within the US, ADM has become an "early target of state AI regulation," giving regulators an overarching concept to hang their policies on (Anderson et al., 2025).

This framing is particularly valuable in addressing challenges stemming from ambiguous definitions and rapid technological evolution. Take the term "artificial intelligence" (AI) as an example. Widely used as an umbrella term, it encompasses a diverse range of technologies (Wang, 2019). The difficulty of pinning down a precise definition is exemplified by the UK government: in its response to a white paper consultation on AI regulation, it explicitly avoided providing a formal definition (Department for Science, Innovation and Technology, 2023; Gallo & Nair, 2023), arguing that the fast-moving nature of the technology makes such a task impractical.

ADM thus provides a flexible and policy-relevant lens for examining the governance of algorithmic systems within the public sector. To contextualise this analysis, the following section outlines the evolving techno-economic landscape that is reshaping regulatory assumptions and priorities.

3. The Dynamic Context for Governance: Techno-Optimism Emerging from Stagnation and Fear

After a half-decade characterised by concern and critique, this research suggests there is a renewed (if still somewhat ambivalent) aura of excitement and techno-optimism surrounding algorithmic technologies. Fuelled by rapid technical advancements, increased investment, and growing deployment, this discourse has disrupted the relatively stable regulatory assumptions of the early 2020s and ushered in a decidedly more innovation-friendly policy climate.

3.1. The AI Backlash (Late 2010s–Early 2020s)

This period, described as "the AI Backlash" (Oremus, 2023), extended the 2010s "techlash" against major technology companies (Atkinson, 2019; Neidig, 2018; Viljoen, 2021) into the realm of algorithmic technologies. Media narratives amplified an "AI technopanic" (Weiss-Blatt et al., 2024), fuelled by open letters emerging from leading scientists calling for pauses on AI production (Future of Life Institute, 2023), and warnings of "extinction level" risks (Roose, 2023). With rising public concern (Pew Research Center, 2023) prompting ambitious regulatory proposals from key global regulators (Biden, 2023; Regulation (EU) 2024/1689, 2024).

Economic indicators mirrored this mood. This period saw the first year-on-year decline in AI funding as global private investing fell 26% from 2021 to 2022 (Stanford Institute for Human-Centered Artificial Intelligence, 2023), with venture funding falling 43% in Q1 2023 (CB Insights, 2023). Uptake and development also encountered issues. AI adoption amongst organisations plateaued in the 50th percentile between 2017 and 2023 (McKinsey & Company, 2022), and developers worried about how a "data wall" would curtail technical evolution ("A.I. companies face," 2024). Moreover, there was a substantial deceleration in the number of AI patent applications (Williams & Sibley, 2022).

3.2. A Recent Resurgence

However, that consensus now looks outdated. Indeed, this article argues that we are at the beginning of a turnaround regarding deployment, funding, technological progress, and, as explored later, regulatory policymaking. According to McKinsey & Company's (2025) global report on the "state of AI" there has been a sharp jump in business deployment of AI, with 78% of surveyed organisations saying they now use AI in at least one business function in 2024, up from 55% on the previous year, marking a major acceleration after the period of prolonged stagnation between 2017 and 2023 (McKinsey & Company, 2025). Accompanying this are extraordinary levels of financial investment: the 2025 *Stanford AI Index* documents that 2024 saw the highest-ever global private investment in AI technologies of over $252 billion (Stanford Institute for Human-Centered Artificial Intelligence, 2025). Such implementation of algorithmic technologies is beginning to result in a notable economic impact, with recent PwC research indicating that AI adoption could boost global GDP by an additional 15 percentage points by 2035, rivalling the growth increment the world began to enjoy during the 19th-century industrialisation (PwC, 2025).

A renewed surge in AI investment and operational adoption is driven by a critical technological evolution—the rise of frontier models capable of methodically "thinking through" problems before solving them (OpenAI, 2024; Pinchai et al., 2025; xAI, 2025). These systems offer significant advantages over earlier generations (Besta et al., 2025), promising greater reliability and interpretability in real-world applications by supposedly emulating nuanced human-like reasoning. Crucially, their rapid emergence is proving disruptive to policy timelines. As European Commission President Ursula von der Leyen acknowledged at the 2025 annual EU budget conference: "[W]hen the current budget was negotiated, we thought AI would only approach human reasoning around 2050. Now we expect this to happen already next year" (Von der Leyen, 2025, para. 3).

3.3. Public Sector Adoption

Importantly, this intensification of AI activity is not confined to the private sector. According to *the G7 Toolkit for Artificial Intelligence in the Public Sector*: "[a]rtificial intelligence is revolutionising how governments work, offering unprecedented opportunities to deliver better public services, improve policy outcomes, enhance public sector productivity, and foster accountability" (OECD & UNESCO, 2024, p.3). Moreover, the Boston Consulting Group estimate "productivity gains of generative AI for the public sector will be valued at $1.75 trillion per year by 2033" (Carrasco et al., 2023, para. 1). These potential advantages have resulted in a substantial increase in algorithmic implementation, strategic policies, and public–private investment initiatives from key global actors, as jurisdictions compete for leadership in AI infrastructure and capabilities (Oxford Insights, 2024).

Nowhere is this adoption more evident than in the US, where the involvement of tech entrepreneurs within the government itself suggests that AI will be a dominant force in both the private and public sectors as the year progresses. Consider, for example, the US Stargate Project (OpenAI, 2025)—a $500 billion AI infrastructure initiative launched in early 2025, described by President Trump as "the largest AI infrastructure project by far in history" (Trump, 2025, para. 7). Accompanying this investment are policy instruments strongly promoting federal adoption of AI technologies. Explored later, these include the Trump Administration's Executive Orders on AI (White House, 2025a) and the accompanying guidance from the Office of Management and Budget (OMB) on AI procurement and use by federal agencies (White House, 2025b).

The US is not alone in this push. Just one month after the Stargate announcement, the EU unveiled its own public–private partnership entitled InvestAI, aiming to mobilise €200 billion to advance AI development and infrastructure across the Union. Commission President Ursula von der Leyen likened it to a "CERN for AI" (European Commission, 2025). Such effort aligns with longstanding EU ambitions to "make the public sector a trailblazer for using AI" (European Commission, 2021, p. 46), a goal that was reaffirmed in the Commission's 2024 study, which "calls for strategic AI adoptions to transform public sector services" (European Commission, 2024a).

The UK government's 2024 *Algorithmic Transparency Report* similarly exemplifies this shift. It frames AI deployment as essential for achieving "technological improvements in critical government services" (Department for Science, Innovation and Technology, 2024, para. 1). Building on this, in January 2025, the government published a plan to drive economic growth and transform the state by scaling up AI across the public sector and acting as a catalyst for private sector AI development (Department for Science, Innovation and Technology, 2025a).

A similar trajectory is also evident in China, where policymakers have long viewed AI as "critical for the future development of innovation, smart industrial systems, and digital life." (Gong & Dorwart, 2024, para. 6). This focus is particularly obvious in the "New Generation Artificial Intelligence Development Plan" (Webster et al., 2017), which emphasises AI's role in boosting economic growth, improving social services, and strengthening national security.

Governments across leading economies are realigning strategically around AI, increasingly framing it as a strategic core asset that can underpin national security and drive economic competitiveness and public sector transformation. The early 2020s "AI Backlash," characterised by declining investment, adoption plateaus, and mounting concerns over limitations and risks, now appears to have been a temporary contraction rather than a structural slowdown.

3.4. Some Concerns Remain

This is not to suggest that recent public sector AI adoption has proceeded without hesitancy. Despite increased deployment and investment as governments position themselves to leverage these technologies, scepticism persists amid increasing acknowledgement of inherent safety and ethical concerns.

Some of the clearest manifestations of this balancing act—between innovation, investment, and safety concerns—emerge in foundational documents for the EU's seminal AI Act. The Commission's original white paper on AI frames the technology as a double-edged sword, noting it "brings both opportunities and risks" and stressing that "while AI can do much good…it can also do harm" (European Commission, 2020a, pp. 9, 11). Similarly, the EU's High-Level Expert Group on AI echoes these concerns, emphasising that AI systems "while bringing substantial benefits to individuals and society…also pose certain risks and may have a negative impact" (European Commission, 2019, p. 2). Such institutional caution is reinforced by non-governmental critics who challenge government AI deployment and harm-mitigation policies (Arda, 2024; Hacker, 2023; Kretschmer et al., 2023).

A broader international perspective emerged in 2025 with the *International AI Safety Report*, prepared by 96 international experts from 30 countries following the Bletchley Park AI Safety Summit (Department for

Science, Innovation and Technology & AI Safety Institute, 2025). It is essentially a cautionary account of AI's capabilities, the associated risks, and possible ways to mitigate them. The concerns raised by these reports have also resonated within academia and the wider AI research community. In February 2025, an open letter signed by AI practitioners and public figures (Milmo, 2025) endorsed an accompanying academic article setting out principles for developing responsible AI (Butlin & Lappas, 2025). Reinforcing these concerns, the *Final Report of the Pissarides Review on the Future of Work and Wellbeing*, published in early 2025 (Institute for the Future of Work, 2025), the latest in a series of cautionary accounts, underscores AI's transformative impact on employment, echoing earlier reports (Department of Education, 2023; Jung & Desikan, 2024).

Against this backdrop of increasing public-sector interest in algorithmic systems, growing institutional caution, and academic concern, mapping the actual state of ADM integration within government also faces notable blind spots. The following section explores these issues.

4. Public Sector Uptake Difficulties

As scholarly assessments remain divided on ADM impacts (Alhosani & Alhashmi, 2024; Contreras & Gil-García, 2024; Mergel et al., 2023), there is a growing number of widely cited instances of ADM failure in the public sector: the UK's A-level exam algorithm (Kippin & Cairney, 2022) and spousal visa AI (Stacey, 2023); US COMPAS sentencing tools (Engel et al., 2024); the Dutch childcare benefit scandal (Hadwick & Lan, 2021); and Australia's "Robodebt" (Chowdhury, 2024), are examples of recurring public-sector risks. However, despite these examples and the mounting interest and investment within public administration, the actual state of ADM deployment in government remains somewhat opaque and difficult to assess comprehensively. Some high-level inter-governmental papers and surveys concentrate upon government AI readiness, such as the OECD's (2024) *Governing with Artificial Intelligence: Are Governments Ready*? and Oxford Insights' (2024) *Government AI Readiness Index*. Certain jurisdictions, such as the EU, have released limited overviews and studies of government use of algorithmic systems (European Commission, 2023, 2024b). However, getting granular detail on system implementation, functionality, or oversight of specific ADM uses within national governments is more difficult. The OECD's November 2024 report on *Algorithmic Transparency in the Public Secto*r highlights information gaps globally. Its mapping of public algorithms repositories found that most public bodies fail to provide "meaningful transparency," or "disclose pertinent and sufficient information to evaluate AI systems" (Gutiérrez & Muñoz-Cadena, 2024, p. 20). Compounding this, systematic under-reporting persists; for example, in the Netherlands, only 5% of reported AI systems are listed in the public registry (Rekenkamer, 2024), and the UK's central algorithmic registry remains sparsely populated years after launch (Booth, 2024b).

The noted OECD report attributes disclosure gaps to institutional incapacity or political reluctance, cybersecurity fears, and IP constraints. There are several cases that would support these assertions, for example, in France, the social security agency Caisse Nationale d'Allocations Familiales initially rejected freedom-of-information requests for its welfare-scoring code, arguing that disclosure would "give fraudsters the keys" (Sénécat, 2023, para. 3). In the Netherlands, indicators used by the SyRI welfare fraud detection algorithm were kept secret until a 2020 court struck down the system, noting its excessive opacity (Zuiderveen Borgesius & van Bekkum, 2021). Across the Atlantic, the US Internal Revenue Service will not release the training or methodology behind its AI-driven audit-selection model (Loricchio & Wallace, 2024).

4.1. UK Case Study

In this context, the UK offers a useful exemplary case study, not least due to its 2025 blueprint "to turbocharge AI," boost public sector adoption, and "make Britain the world leader" in these technologies (Department for Science, Innovation and Technology, 2025e). Yet this ambition sits in tension with the limited public data on public ADM and the recurring challenges associated with implementation.

Historically, the UK has been a leader in governmental use of digital technologies, ranking first in the UN's (2016) e-government development index and third in the OECD's (2023) *Digital Government Index*. However, an up-to-date picture, including the range of deployment of algorithmic technologies, is more difficult to find. Following a consultation on the last *Government's National Data Strategy* (Department for Science, Innovation and Technology, 2024), a mandatory Algorithmic Transparency Recording Standard now exists. However, so far, there is only a small repository of fairly basic information (Cabinet Office, 2025). There is a clear desire to develop such technologies: the Blueprint for Modern Digital Government (Department for Science, Innovation and Technology, 2025c, p. 4) which accompanies the State of Digital Government Review (Department for Science, Innovation and Technology, 2025b), recognises the potential to "catalyse a wholesale reshaping of the public sector." However, there is also a clear recognition of the scale of the challenges that face the newly established Government Digital Service (Department for Science, Innovation and Technology, 2025c, p. 1). Perhaps as a result, a National Audit Office report from 2024 suggests a relatively small and slow uptake of AI across government, with only 37% of the 87 government bodies responding to the National Audit Office survey reporting that they deploy AI, although it is acknowledged that some 70% are piloting or planning AI systems (National Audit Office, 2024).

Financial and logistical issues impact this slow uptake. But there are also regulatory concerns. As the National Audit Report acknowledges: "Government standards and guidance to support responsible and safe adoption of AI are still under development" (National Audit Office, 2024, p. 10). Those that do exist tend to be rather general in their advice; for example, the current action plan (Department for Science, Innovation and Technology, 2025a) is largely focused upon development, with a desire for "pro-innovation" regulation. Meanwhile, the *Generative AI Framework for HM Government* (Cabinet Office, 2024) can offer only 10 principles in line with very general ideas of lawfulness and ethics, while the AI Safety Institute offers a five-step process-driven approach (AI Safety Institute, 2025). More specific policies, such as Lord Clement-Jones' private members bill on Public Authority Algorithmic and Automated Decision-Making Systems Bill (House of Lords, 2024), advance slowly.

The UK is an illustrative example of the wider international challenges posed by the current regulatory landscape and underutilisation of data repositories. Effectively, it remains difficult to determine exactly where and how algorithmic technologies are being deployed within the public sector, and, as a result, what the most pressing risks are. Amid this ambiguity, this article analyses one area of the public sector where ADM has gained major global traction: the court system. This setting offers a valuable lens through which to examine the benefits, challenges, and risks that government agencies must navigate when implementing ADM systems. Serving as a microcosm of the broader challenges surrounding ADM adoption in the public sector, particularly in the noted absence of comprehensive information on other cases of use.

4.2. AI in the Courts: An Exemplar

Courts have historically been early adopters of new technologies, in part because the legal system (especially judging) has long attracted the interest of technology specialists. This may arise from a common, though reductive, perspective among computer scientists, who often view law as a more sophisticated version of games like chess or Go, where computational power is presumed to surpass human judgment. However, as Morison and Harkens (2019) point out, law is not merely about applying the "correct" rule to a given set of "facts"; it is a far more complex and socially embedded process.

Nonetheless, there is often considered to be potential for the development of AI in the court context, particularly given the promise of increased efficiency. In England and Wales alone, some 3.1 million cases pass through the courts annually. An AI system working continuously at speed might be seen to improve access to justice and reduce costs—arguments that perhaps underpin much of the UK government's support for such algorithmic systems in the judicial context (Mark et al., 2024). This trend is not unique to the UK. In the US, ADM tools such as COMPAS (Brennan & Dieterich, 2017) and Advancing Pretrial Policy & Research (APPR, 2025) have already been incorporated into judicial processes. Similarly, in Europe, models like OxRec have been trialled, with the European Commission's communication *Digitisation of Justice in the European Union: A Toolbox of Opportunities* stating: "AI applications can bring a lot of benefits" (European Commission, 2020c, p. 10). Perhaps the most ambitious initiative can be found in China, where the Supreme People's Court has issued guidance requiring courts to develop competent AI systems by 2025 (Sourdin, 2021; Supreme People's Court, 2022; Xia, 2024).

Within the judiciary itself, despite warnings that ADM systems may entrench existing data biases, undermine accountability, and erode meaningful oversight (Big Brother Watch, 2023; Public Law Project, 2023), it is perhaps relatively uncontroversial to suggest that some judicial processes are amenable to a degree of automation or assistance, with some judicial actors appearing open to such assistance. For example, Lord Justice Birss famously commented that he has used ChatGPT and found that it can be "jolly useful" in giving a summary of an area of law which he knew already (Corfield, 2023). Others go even further, using LLMs more directly in producing judgments (Digital Watch Observatory, 2023; Gutiérrez, 2024; Taylor, 2023). More ambitiously, some proponents argue that ADM tools may even be an improvement on human judging, minimising bias and operating more objectively, particularly in routine administrative decisions that follow procedural steps, such as parking fines or license applications (Alessa, 2022; Katsh & Rabinovich-Einy, 2017). However, matters involving criminal sentencing or child custody may be thought to require a more personalised, human approach. Clearly, there is a spectrum of legal decisions, from what might be seen as purely "administrative" decisions (where automation may be unproblematic) to the more obviously "judicial" (which may seem to require the input of a human; Morison & McInerney, 2025). As such, despite the promises of reduced cost and potential efficiency gains, the wholesale replacement of judges by ADM remains unlikely, but clearly, there is room and appetite for some algorithmic tools (Ministry of Justice, 2025).

Regulatory activity around ADM in the judiciary remains relatively limited, reflecting broader trends in public sector adoption of algorithmic technologies. In the EU, the 2018 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment outlines five general principles, but it predates recent advances in generative AI. The UK has issued only brief guidance, offering high-level risk categories and

placing responsibilities on judges for all material produced (Courts and Tribunals Judiciary, 2023). In the US, judicial engagement is emerging through case law, for example, Ross v United States (2025) acknowledges AI's utility but warns against delegating decision-making. China, despite its rapid deployment of AI in courts, has given comparatively little attention to regulatory safeguards. By contrast, Australia offers a more proactive model, with a collaborative initiative between judges and academics aimed at monitoring AI developments and creating practical judicial guidance (Australian Institute of Judicial Administration, 2023).

Intriguingly, there appears to be a clear desire amongst practitioners for further guidance. The UNESCO Global Judges' Initiative (UNESCO, 2024) surveyed judicial understanding and use of AI, with 72% of respondents believing there should be mandatory rules for judges in this area (UNESCO, 2024, p. 11).

This snapshot of how ADM has been approached in the judicial context exposes the core tensions shaping its use across the public sector. It puts a focus on decision-making and stresses the values of publicness. The main issues—balancing efficiency with fairness, and technological progress with meaningful human judgment—are mirrored across other government contexts and highlight a central balancing act in the use of ADM. States increasingly prioritise algorithmic-driven modernisation, yet must reconcile these ambitions with accountability gaps where transparency and human oversight are missing. Crucially, how these tensions are understood as "problems" directly shapes governance. Therefore, the following sections analyse how key jurisdictions problematise such tensions, examining emerging policies and regulatory discourse.

5. Problematization

While documenting the resurgent techno-optimism, fuelled by private-sector innovation, public-sector AI adoption (e.g., judicial systems), and state ambitions, this analysis reveals fundamental tensions in governing priorities. Despite apparent enthusiasm for the transformative potential of algorithmic technologies, each jurisdiction constructs the associated risks and responsibilities in distinct ways. Crucially, jurisdictions diverge radically in framing algorithmic "problems": the US embraces light-touch governance versus the EU's precautionary paradigm, for example. To interrogate this divergence, we apply aspects of Bacchi's (2009) WPR framework. This approach examines how phenomena become constituted as governance "problems" (Foucault, 1985, p. 115) and why solutions reflect specific ideological and institutional positionalities (Miller & Rose, 2008).

While the following sections do not apply Bacchi's WPR framework in a step-by-step manner, its core analytical questions serve to guide the comparative discussion that follows. In particular, attention is paid to how ADM is problematized in each jurisdiction, what assumptions and rationalities underpin these framings, and what silences are produced. By doing so, the analysis moves beyond descriptive comparison to interrogate the underlying political rationalities, institutional logics, and techno-economic imperatives shaping governance strategies in the US, EU, and China. This approach allows for a deeper reflection on how problem representations influence governance and the broader dynamics of power and responsibility, particularly when algorithmic systems take on tasks, such as judicial or administrative decision-making, once performed exclusively by human experts.

As noted, this analysis focuses on recent domestic regulation and national policy discourse. In the context of AI regulation, Veale et al. (2023) identify six regulatory modalities, including ethical codes, industry

governance, licensing, standards, and international agreements, but highlight converging and extraterritorial domestic legislation as the most concrete. Whilst acknowledging the relevance of other approaches, this research centres on that last modality, focusing on how jurisdictions assert authority and shape norms through national domestic regulatory frameworks. This focus offers a grounded lens through which to analyse the problematization and governance of ADM systems.

6. Key Players

Focusing on the EU, US, and China, described by Bradford (2023) as competing "digital empires," this account sketches out their contest to shape global AI governance. Bradford maintains that because of the global nature of the digital economy, these leading regulatory models extend across jurisdictions, impacting foreign societies. Other nations are almost forced to align with one of the three models expounded by the US, EU, and China. This triadic dynamic is further illuminated by Schneider's (2025) work on digital sovereignty, which positions the EU's rights-based framework as a "third way" between the US market-driven approach and China's state-centric techno-authoritarianism. Given the noted absence of comprehensive sector-specific legislation governing ADM itself, this article concentrates upon emerging regulatory policies and related discourse, developed under the broader frameworks of AI or digital governance in general, which function as proxies through which the governance of ADM can be analysed.

6.1. United States

In the US, regulation of algorithmic technologies is notably fragmented, relying on a combination of existing federal laws, non-binding guidance, and state-level initiatives rather than a single comprehensive framework. Moreover, the overarching direction set by the executive branch tends to fluctuate with the political orientation of the incumbent administration, producing distinct shifts between Republican and Democratic leadership. Yet, across these shifts, certain commonalities remain that speak to a uniquely American way of framing algorithmic problems.

For instance, recent executive initiatives, most notably the Removing Barriers to American Leadership in Artificial Intelligence Executive Order (EO 14179) and the July 2025 AI Action plan (White House, 2025c), represent a longstanding commitment to fostering an innovation-friendly environment and reflect a deep-seated political rationality of permissionless innovation (Thierer, 2016), in which regulation itself is often portrayed as a problem. This orientation draws on longstanding US traditions of free-market liberalism and deregulation, particularly in the governance of emerging technologies. A formative example is the Clinton Administration's 1997 Framework for Global Electronic Commerce. This laid the foundations for internet governance. Its first two principles, "the private sector should lead" and "governments should avoid undue restrictions on electronic commerce," rejected prior regulatory models, embracing self-regulation, enforced through market competition and tort law (White House, 1997, p. 3). Technology was constructed as a domain of market freedom rather than state control, an understanding that continues in contemporary algorithmic regulation today.

Consider, for example, the recent AI Action plan, which requires the removal of any policy "that unnecessarily hinder[s] AI development or deployment" (White House, 2025e, p. 3), a clear extension of the earlier executive order 14179 and its revocation of any "existing AI policies and directives that act as barriers to American AI

innovation" (White House, 2025d, s. 1). This deregulatory framing was further evident in the specific repeal of President Biden's expansive executive order 14110 on "safe, secure, and trustworthy development and use of artificial intelligence" (Biden, 2023). Executive order 14110 had imposed substantive obligations on federal agencies to ensure transparency, safeguard against algorithmic discrimination in public services, and manage risks in government use of AI. It was arguably the most far-reaching executive action on algorithmic governance to date. Its revocation also raises uncertainty over the future of the underlying Blueprint for an AI Bill of Rights. This was an earlier Biden-era document that acted as a "national values statement," laying out principles for safe and effective algorithmic systems, stating that citizens "should not face discrimination by algorithms" and could seek review of consequential automated decisions (Office of Science and Technology Policy, 2022, para. 6).

Further insight can be garnered from memoranda emerging from the OMB. Mandated by the AI in Government Act of 2020, these documents now constitute the de facto regulatory framework for AI systems across the US federal government. Executive Order 14179 specifically revised OMB's guidance, eliminating all references to the Blueprint for an AI Bill of Rights previously embedded in the Biden-era M-24-10/M-24-18. This revision signals a reprioritisation of objectives, with the opening directives of OMB Memorandum 25-21 stating: "Agencies must remove barriers to innovation and provide the best value for the taxpayer," "Agencies must empower AI leaders to accelerate responsible AI adoption," and "Agencies must ensure their use of AI works for the American people" (Office of Management and Budget, 2025, p2). Innovation now explicitly precedes safeguards in the regulatory hierarchy.

Effectively, the problem is currently framed as "burdensome requirements" stifling innovation (White House, 2025a), with the current administration responding with deregulatory guidance. However, it should be noted that whilst this deregulatory attitude has perhaps accelerated under the new presidency, it is reflective of historical US policymaking and the underlying political rationality of permissionless innovation. Even under the more guardrail friendly Biden administration, commitments to algorithmic safety and security were largely realised through soft law instruments, revealing a persistent reluctance to impose binding regulatory constraints on algorithmic innovation.

Indeed, there is a notable institutional logic commonality across administrations, an apparent desire to avoid new binding regulation. President Trump's first term set the tone with executive order 13859 (Trump, 2019) and the American AI Initiative, which instructed agencies to "avoid regulatory over-reach" (Office of Science and Technology Policy, 2020, p. 15) and to promote innovation within existing statutory mandates. Later, OMB Memorandum M-21-06 crystallised that approach, advising agencies that they should "consider either not taking any action" or adopt "non-regulatory approaches" when regulating these technologies (OMB, 2020). President Biden arguably preserved this preference with the Blueprint for an AI Bill of Rights (Office of Science and Technology Policy, 2022), the voluntary NIST AI-RMF (National Institute of Standards and Technology, 2023), and a range of federal agencies' voluntary guidelines and frameworks, all resting on self-assessment and voluntary commitments rather than enforceable mandates.

Consequently, individuals seeking redress for algorithmic decisions causing harm must generally rely upon existing legal frameworks and established federal agencies. For instance, the Fair Credit Reporting Act, the Civil Rights Act, and the Equal Protection Clause of the Fourteenth Amendment have been invoked in the past to challenge discriminatory outcomes produced by automated systems. Meanwhile, a joint statement from

four federal agencies, the Consumer Financial Protection Bureau, Department of Justice, Equal Employment Opportunity Commission, and Federal Trade Commission (2023) makes it clear that they will apply current sector specific legal policies to algorithmic harms, stating that "existing legal authorities apply to the use of automated systems and innovative new technologies just as they apply to other practices." However, emerging from this logic is a relative silence about rights or issues that fall outside the traditional US legal framework. For example, the right to an explanation of algorithmic decisions, a topic of discussion in the EU, is not explicitly recognised in most US regulatory policies; the Blueprint for an AI Bill of Rights hints at it but doesn't actively enforce it. Similarly, protections against more diffuse harms are not developed in US discourse.

Intriguingly, reflective of this constitutional minimalism logic, regulatory attempts to provide broader and more direct governance often do not prosper. For instance, the Algorithmic Accountability Act of 2022 (H.R. 6580), introduced in the US House of Representatives, sought to mandate impact assessments for ADM systems to mitigate potential biases. Yet, the bill stalled after being referred to the Subcommittee on Consumer Protection and Commerce and ultimately failed to pass before the 117th Congress adjourned. Such actions reveal a further representation within US policy that existing legal norms and established regulators are presumed to provide adequate cover for algorithmic harms.

It should be noted that there is perhaps more of a push for regulation at the state level, with various narrow legislative policies on AI emerging in recent years (California State Legislature, 2024; Colorado General Assembly, 2024; New York State Legislature, 2024; Virginia General Assembly, 2025). However, these state-level initiatives often face similar obstacles to those encountered at the federal level. Consider, for example, California's SB1047, which would have implemented some algorithmic safeguards on frontier models. Despite some support for this bill (Lovely, 2024; Nazzaro, 2024), opponents pushed back (Abbott, 2024; Bensinger, 2024), arguing that strict regulatory thresholds would hinder competitiveness and fail to capture the full scope of risks. As a result, the bill was vetoed by Governor Gavin Newsom in September 2024, arguably reflecting a broader, recurring pattern in US policymaking, where concerns surrounding technological leadership and stifling innovation often outweigh a call for precautionary safeguards.

Moreover, it should also be acknowledged that there is a reluctance at the executive level to allow a fragmented regulatory approach across states, with OMB guidance encouraging federal agencies to consider pre-empting state laws. More recently, a failed provision placed into a draft federal budget bill attempted to ban all state regulation of "artificial intelligence models, artificial intelligence systems, or automated decision systems' for the next ten years" (US House of Representatives, 2025). Whilst absent in the final bill, such a provision perhaps reveals a determination by some policymakers in the US to ensure that state regulation does not become a barrier to innovation and investment.

In sum, ADM governance in the US is structured in part by administration cycles, but underlying this is a durable understanding of pro-innovation political rationalities, soft-law institutional logics, and techno-economic imperatives prioritising growth dynamics that arguably also create certain blind-spots around unique algorithmic harms. In this context, innovation often outweighs other concerns for this key actor, with regulatory policy and discourse framing an "innovation gap" as a central problem. Moreover, the recent increase of deregulatory action under the current administration is perhaps the clearest illustration of the broader trend mentioned in Section 3 of this article, where regulatory design is progressively facilitating technological leadership and the pursuit of innovation over a desire for protective guardrails.

6.2. European Union

The EU's approach to ADM presents a near antithesis to the innovation-led US model. Fundamentally precautionary in nature, it reflects a conviction that emerging technologies risk harming foundational concepts such as fairness, privacy, and fundamental rights, necessitating new and comprehensive regulatory architecture. This stance embodies Europe's longstanding "regulatory state" tradition (Majone, 1994) and implicitly acknowledges the inadequacy of existing legal safeguards.

This framing reflects a deeper political rationality rooted in the EU's long-standing commitment to the precautionary principle. This is a key governing approach in European regulation, originating in international environmental law and incorporated into European governance through the Maastricht Treaty 1992 and Article 191 of the Treaty on the Functioning of the European Union. It mandates that policymakers "err on the side of caution by adopting relatively stringent regulations even in cases where the scientific evidence...is unclear, inconclusive, ambiguous or uncertain" (Vogel, 2012). Although traditionally applied in areas such as environmental protection and public health, the European Commission has acknowledged that "its scope is much wider" (European Commission, 2000). Indeed, scholars suggest that the principle is increasingly shaping digital regulation, as reflected in the risk-based response of the AI Act (Howell, 2023).

The EU, therefore, views legal intervention not as a constraint on innovation, but as a precondition for societal trust and technological legitimacy, foundations considered central to the uptake and integration of AI across the single market. In this way, regulation functions as a techno-economic imperative in its own right, a necessary condition for driving long-term innovation and growth.

Central to this framing is the concept of "trustworthiness." The European Commission has expressly stated that "trustworthiness is seen as a crucial feature of European AI" and central to its vision of "human-centric" AI (European Commission, 2020b, p. 4). This "foundational ambition" (European Commission, 2019) is defined by the EU High Level Expert Group on AI as "AI that is legally compliant, ethically adherent, and socio-technically robust" (European Commission, 2020b, p. 3), reflecting aspects of the EU charter of fundamental rights. This normative ambition supports both a political and institutional logic, a rights-based regulatory culture in which public trust is to be constructed through formal legal design.

The AI Act epitomises this regulatory philosophy. Its stated purpose, "to improve the functioning of the internal market and promote the uptake of human-centric and trustworthy AI, while ensuring a high level of protection of health, safety, and fundamental rights..." (Regulation (EU) 2024/1689, 2024, Article 1), operationalises precaution through a risk-based framework. This imposes technical safeguards, reflecting the EU's product-safety logic, effectively extending the New Legislative Framework model into the algorithmic domain. Notably, Annex III designates public sector systems affecting essential services or rights (e.g., welfare eligibility) as "high-risk" (Regulation (EU) 2024/1689, 2024, Annex III), directly subjecting state algorithmic decision-making to stringent oversight. This architecture reflects the EU's core problem framing. Here, trust is not presumed but rather constructed through regulatory compliance, and an underlying techno-regulatory optimism assumes that algorithmically induced harms can be largely mitigated, and trust engendered through pre-emptive regulatory demands.

Within this Act, and of relevance to this article, is the importance placed upon decision-making, with Article 6 and Recital 53 (Regulation (EU) 2024/1689, 2024) underscoring the centrality of the decision-making function in assessing the risk level of a system. Moreover, rather than supplanting previous regulations, the act builds upon protections found in the seminal GDPR, including Article 22, which prohibits certain decisions from being made solely by automated means (Regulation (EU) 2016/679, 2016, Art. 22). The GDPR also guarantees individuals the right to meaningful information about the logic involved in automated processing, often interpreted as a "right to explanation," Articles 13–15. There has been extensive academic and practical debate around this, (Edwards & Veale, 2018; Kaminski, 2018; Wachter et al., 2017) which the AI Act helps clarify, confirming that with high-risk systems "an affected person now has a right to meaningful explanations on the role of the AI system in the decision-making and the main elements of the decision made" (Regulation (EU) 2024/1689, 2024, Article 86). Consequently, the EU's regulatory approach to decisions made by algorithms combines a broad data protection-based restriction on fully automated decisions (GDPR, Article 22) with more targeted oversight through the AI Act's risk-based framework, again pointing to a techno-regulatory optimism. Further, this also reveals a regulatory silence, the assumption that technical transparency equates to substantive accountability, when, in practice, explanation rights often fail to address underlying power asymmetries or structural inequities.

Yet despite its rights-forward posture, recent developments reveal growing tension between the AI Act's normative ambitions and economic pragmatism. Reports suggest efforts to dilute certain provisions for a more innovation-friendly environment (Espinoza & Dubois, 2025) while techno-regulation optimists warn they are losing the narrative battle. As UN AI Advisory Board co-chair Carme Artigas contends, European companies believe the "absolute lie" that the AI Act is killing innovation (Greenacre, 2024). This friction materialised during the AI Act's late-stage negotiations, where the major economies of France, Germany, and Italy pushed to dilute the regulatory burdens for industry (Perrigo, 2023). The pressure persists post-adoption, with the European Commission VP for digital policy Henna Virkkunen recently emphasising the need to avoid "creating more reporting obligations for our companies," while cutting "red tape" (Foy & Moens, 2025). This shift signals the battle over competing techno-economic imperatives within the EU's governance logic, where competitiveness and regulatory burden are now being weighed more heavily. Perhaps "the world's regulatory superpower" (Malloy, 2023), which has often "led the charge on digital regulation" (Hobbs, 2020), is not immune to the sway of innovation and investment at the price of regulatory accountability for ADM. Consider, for example, the quiet withdrawal of the EU's AI Liability Directive, just days after the private-public €200-billion InvestAI fund was announced in February 2025.

To summarise, the EU's governance of ADM is shaped by a precautionary political rationality rooted in rights-protection and public trust, an institutional logic grounded in layered risk-based legal oversight, and a techno-economic imperative that seeks to align innovation with regulation. The overarching framing of a "trust deficit" positions ADM as a domain requiring active regulation to safeguard rights and maintain public confidence. However, recent discourse around scaling back regulatory obligations reveals a growing tension within this model, one that perhaps reflects the broader trend noted above, where the EU's commitment to regulation is increasingly tempered by global investment pressures and the perceived need to avoid excessively burdensome regulation.

6.3. China

Since the release of the Next Generation AI Development Plan in 2017, authorities have consistently emphasised AI's economic potential and strategic value, with the stated aim of becoming a world leader in AI by 2030. This agenda includes public sector transformation, with a concerted push to use algorithmic technologies to increase efficiency, exemplified by the Supreme People's Court's directives to integrate AI into judicial processes, as discussed in Section 4. In this context, China's emerging regulatory framework for ADM in the public sector is shaped by the state's broader techno-economic imperative to maximise the economic and strategic benefits of algorithmic technologies through indigenousness innovation and centralised technological control.

A distinguishing feature of China's approach is a clear emphasis on ensuring that the deployment of algorithmic technologies does not pose risks to political and social stability (Gong & Dorwart, 2024; National Technical Committee, 2024; Sheehan, 2023). Algorithms are seen as economically necessary tools to improve services, enforce laws, and promote trust in society, but they can also create stability risks if not properly controlled, with official discourse warning of dangers such as the spread of misinformation or destabilising public opinion. As a result, many commentators argue that China's regulatory approach is predominantly driven by state interests rather than individual rights, as exemplified by initiatives like the widely criticised social credit scoring system (Boyer, 2022; Chin & Lin, 2022; Roberts et al., 2021; US Department of Defense, 2018; Zeng, 2020). In this logic, the "problem" is not necessarily the erosion of individual freedoms but the state's diminished capacity to maintain control. Regulation, therefore, often centres on ex-ante content control and continuous monitoring rather than ex-post mechanisms for individual redress. This approach is embedded in the Xi-era political rationality of "holistic (overall) national security" (总体国家安全观), which, as Blanchette (2020, para. 2) notes, "has come to subsume nearly all elements of policymaking and political considerations." Under this framework, security is redefined as an all-encompassing imperative that includes political, cultural, and societal stability, elevating ideological conformity and social cohesion to matters of national security, including those arising from emerging technologies (Liao, 2025).

In terms of specific governance, in the early 2020s, China's regulatory landscape for ADM moved to a more formal oversight. This period is sometimes referred to as a regulatory crackdown (Cao, 2025; Hsu, 2021). A key milestone at this time was the Personal Information Protection Law (Creemers & Webster, 2021), widely regarded as China's GDPR-like data privacy law, and notably one of the first policies to regulate ADM. It mandates that automated decisions derived from personal data be carried out transparently, fairly, and without discrimination, and grants individuals the right to an explanation and the ability to refuse certain automated decisions (Creemers & Webster, 2021). However, it contains broad exemptions for state agencies, arguably weakening protections for individuals and reinforcing a problem representation that concentrates upon protecting state interests. Indeed, the Personal Information Protection Law reflects an intriguing silence found throughout Chinese regulatory policies, which predominantly problematize algorithmic systems deployed by private platforms, casting the state as either a neutral regulator or a trustworthy user of ADM. Risks associated with state-led uses, such as those in surveillance, policing, or the Social Credit System, receive little scrutiny. As such, issues that are prominent in Western AI debates, such as due process for automated decisions or independent oversight, receive far less attention in Chinese discourse.

Reinforcing the above arguments, many Chinese policies contain a recurring requirement for algorithmic technologies to uphold "core socialist values" (Central Committee of the Communist Party of China, 2013), ensuring they do not generate outcomes that contradict the ideology of the Communist Party (Ye, 2023). For example, the Cyberspace Administration of China adopted mandatory provisions like the Internet Information Service Algorithmic Recommendation Management Provisions, which primarily addressed content curation algorithms, requiring aspects of transparency, user opt-out provisions, alongside the cultivation of "positive energy" (Cyberspace Administration of China, 2021, Art. 6), in other words, alignment with the party-state's ideological priorities. The Cyberspace Administration of China later followed with Measures for the Management of Generative AI Services (Cyberspace Administration of China, 2023), establishing content restrictions and security assessments for any public-facing generative AI tools, obligating outputs to adhere to core socialist values.

More recently, in 2024, the National Information Security Standardization Technical Committee (TC260) released its AI Safety Governance Framework, which draws upon previously established principles, including fairness, transparency, and the safeguarding of core socialist values. Whilst not a law, this framework is still influential and will likely impact any future regulations. Sheehan (2023) notes the outcome-based requirements found in these policies, such as the requirement that content again reflect socialist core values. This logic is further apparent in the 2025 AI labelling rules, which mandate both visible and machine-readable labelling of created content, to put an end to the misuse of AI generated technologies and the spread of false information (Allen & Gledhill, 2025, para. 2).

These developments illustrate the unique way ADM is problematized within China's governance of this arena, shaped by a combination of state-centric political rationalities, pre-emptive institutional logics, and techno-economic imperatives that prioritise both ideological cohesion and technological leadership. Political stability and party authority are treated as central regulatory concerns, reframing the "problem" of ADM as a "stability risk," whether that be to social harmony, core socialist values, or national security. Meanwhile, China's techno-economic ambitions position algorithmic systems as essential to domestic modernisation and global competitiveness, embedding AI into public-sector transformation while tightly managing its risks. These dynamics produce a governance model that permits some individual rights, such as transparency and contestability, but does so selectively and often with broad exemptions for state actors.

Interestingly, in what may be a response to a slowing economy, growing pressure to compete with the United States, and a recognition of the perceived economic benefits of algorithmic technologies, China has notably softened elements of its regulatory discourse since 2022. Indeed, Singer and Sheehan (2025, p. 2) note distinct regulatory phases that correspond broadly with the shifts described in Section 3 of this research, a "restrictive 'Crackdown Era' (2020–late 2022), when the CCP reasserted control over tech companies," followed by a more "pragmatic 'Catch-Up Era' (2022–early 2025), that loosened restrictions to boost economic growth." This latter phase is perhaps best exemplified by the 2023 Interim Measures for Generative AI Services, which were noticeably lighter than the original draft following industry pushback (Sun & Zeng, 2024). Notably, Article 3 of the final version commits regulators to a policy of "tolerant and cautious graded management" that seeks to "encourage innovation" (Cyberspace Administration of China, 2023, Art. 3).

7. Conclusion

As this Bacchi-inspired analysis reveals, divergent problem framings drive distinct ADM governance approaches across jurisdictions as they engage with increasing use of this technology in the public sector. The US innovation gap framing supports a deregulatory agenda that prioritises market-led innovation. The EU emphasises a trust deficit that warrants stronger regulation to protect fundamental rights and garner trust. Meanwhile, China's repeated desire to maintain social stability and core socialist values results in ADM as a stability risk, requiring some unique outcome-focused regulatory requirements.

These digital empires channel common safety challenges through nationally distinct governance pathways, whether through existing law and federal agencies in the US, new EU rights-based legislation, or China's collectivist regulatory mechanisms. Yet, concurrently, a convergent regulatory softening emerges. This convergence risks a governance paradox: even as states recognise AI's novel risks, their policy actions replicate the very "move fast and break things" ethos that contributed to the 2020s backlash, potentially sacrificing long-term public trust for short-term gains.

The regulatory softening and diminishing post-2020s accountability mechanisms are likely driven by renewed techno-optimism, economic competitiveness pressures, national security imperatives, and/or global leadership ambitions. While most visible in US deregulatory actions, parallel shifts emerge in EU discourse and China's innovation-prioritising actions.

Crucially, despite divergent problem framings (innovation gap vs. trust deficit vs. stability risk), all three digital empires increasingly privilege innovation over accountability, lured by service transformation promises and the elusive prize of sustained economic growth.

References

Abbott, A. (2024). *Coalition letter opposing California SB 1047*. Law & Economics Center. https://laweconcenter. org/resources/coalition-letter-opposing-california-sb-1047

Advancing Pretrial Policy & Research. (2025). *About the public safety assessment*. https://advancingpretrial.org/ psa/factors

A.I. companies face new restrictions on data use. (2024, July 19). *The New York Times*. https://www.nytimes. com/2024/07/19/technology/ai-data-restrictions.html

AI Safety Institute. (2025). *Principles for safeguard evaluation*. https://www.aisi.gov.uk/work/principles-for-safeguard-evaluation

Alessa, H. (2022). The role of artificial intelligence in online dispute resolution: A brief and critical overview. *Information & Communications Technology Law, 31*(3), 319–334.

Alhosani, K., & Alhashmi, S. (2024). Opportunities, challenges, and benefits of AI innovation in government services: A review. *Discover Artificial Intelligence, 4*, Article 18.

Allen & Gledhill. (2025, April 2). *China further regulates AI-generated content*. https://www.allenandgledhill. com/cn/publication/articles/30284/further-regulates-ai-generated-content

Anderson, H., Reem, N., & Susas, J. (2025, March 7). *Automated decision-making emerges as an early target of state AI regulation*. White & Case LLP. https://www.whitecase.com/insight-alert/automated-decision-making-emerges-early-target-state-ai-regulation

Arda, O. (2024). The European Union's AI Act: Navigating exemptions and ensuring comprehensive safety. *Journal of European Technology Law, 15*(2), 45–67.

Atkinson, R. (2019). *A policymaker's guide to the "techlash"—What it is and why it's a threat to growth and progress*. Information Technology & Innovation Foundation. https://itif.org/publications/2019/10/28/policymakers-guide-techlash

Australian Institute of Judicial Administration. (2023). *AI decision-making and the courts: A guide for judges, tribunal members and court administrators*. https://ssrn.com/abstract=4162985

Bacchi, C. (2009). *Analysing policy: What's the problem represented to be?* Pearson Education.

Bensinger, G. (2024, August 21). Big Tech wants AI to be regulated. Why do they oppose a California AI bill? *Reuters*. https://www.reuters.com/technology/artificial-intelligence/big-tech-wants-ai-be-regulated-why-do-they-oppose-california-ai-bill-2024-08-21

Besta, M., Barth, J., Schreiber, E., Kubicek, A., Catarino, A., Gerstenberger, R., Nyczyk, P., Iff, P., Li, Y., Houliston, S., Sternal, T., Copik, M., Kwaśniewski, G., Müller, J., Flis, Ł., Eberhard, H., Niewiadomski, H., & Hoefler, T. (2025). *Reasoning language models: A blueprint*. arXiv. https://arxiv.org/pdf/2501.11223v3

Biden, J. R. (2023). *Safe, secure, and trustworthy development and use of artificial intelligence* (Executive Order 14110). Federal Register. https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Big Brother Watch. (2023). *Big Brother Watch's response to the government's consultation the "A pro-innovation approach to AI regulation" white paper*. https://bigbrotherwatch.org.uk/wp-content/uploads/2023/06/Big-Brother-Watch-response-to-Govt-White-Paper-on-AI.pdf

Blanchette, J. (2020). *Ideological security as national security*. Center for Strategic & International Studies. https://www.csis.org/analysis/ideological-security-national-security

Booth, R. (2024a, November 11). AI tool could influence Home Office immigration decisions, critics say. *The Guardian*. https://www.theguardian.com/uk-news/2024/nov/11/ai-tool-could-influence-home-office-immigration-decisions-critics-say

Booth, R. (2024b, November 28). UK government failing to list use of AI on mandatory register. *The Guardian*. https://www.theguardian.com/technology/2024/nov/28/uk-government-failing-to-list-use-of-ai-on-mandatory-register

Boyer, R. (2022). Platform capitalism: A socio-economic analysis. *Socio-Economic Review, 20*(4), 1857–1885.

Bradford, A. (2023). *Digital empires: The global battle to regulate technology*. Oxford University Press.

Brennan, T., & Dieterich, W. (2017). Correctional offender management profiles for alternative sanctions (COMPAS). In F. E. Cullen, P. Wilcox, & J. M. Lux (Eds.), *Handbook of recidivism risk/needs assessment tools* (pp. 49–75). Wiley. https://doi.org/10.1002/9781119184256.ch3

Butlin, P., & Lappas, T. (2025). *Principles for responsible AI consciousness research*. ResearchGate. https://www.researchgate.net/publication/387975909_Principles_for_Responsible_AI_Consciousness_Research

Cabinet Office. (2024). *Generative AI framework for HMG*. UK Government. https://www.gov.uk/government/publications/generative-ai-framework-for-hmg

Cabinet Office. (2025). *Find out how algorithmic tools are used in public organisations*. UK Government. https://www.gov.uk/algorithmic-transparency-records

California State Legislature. (2024). *Safe and secure innovation for frontier artificial intelligence models act* (Senate Bill 1047).

Cao, A. (2025, January 27). China's Weibo adjusts algorithms to improve 'public values' amid government crackdown. *South China Morning Post*. https://www.scmp.com/tech/big-tech/article/3296450/chinas-weibo-adjusts-algorithms-improve-public-values-amid-government-crackdown

Carrasco, M., Habib, C., Felden, F., Sargeant, R., Mills, S., Shenton, S., Ingram, J., & Dando, G. (2023). *Generative AI for the public sector: From opportunities to value*. BCG. https://www.bcg.com/publications/2023/unlocking-genai-opportunities-in-the-government

CB Insights. (2023). *State of AI Q1'23 report*. https://s3-us-west-2.amazonaws.com/cbi-content/reports/CB-Insights_Artificial-Intelligence-Report-Q1-2023.pdf

Central Committee of the Communist Party of China. (2013, December 23). *Decision of the Central Committee of the Communist Party of China on some major issues concerning comprehensively deepening reform*. Xinhua. http://www.gov.cn/jrzg/2013-11/15/content_2528179.htm

Chin, J., & Lin, L. (2022). *Surveillance state: Inside China's quest to launch a new era of social control*. St. Martin's Publishing Group.

Chowdhury, S. (2024). Technology is never neutral: Robodebt and a human rights analysis of automated decision-making on welfare recipients. *Australian Journal of Human Rights*, *30*(1), 20–40. https://doi.org/10.1080/1323238X.2024.2409620

Colorado General Assembly. (2024). *Colorado artificial intelligence act* (Senate Bill 24-205).

Consumer Financial Protection Bureau, Department of Justice, Equal Employment Opportunity Commission, & Federal Trade Commission. (2023). *Joint statement on enforcement efforts against discrimination and bias in automated systems*. https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf

Contreras, R., & Gíl-García, J. (2024). Exploring the negative impacts of artificial intelligence in government: The dark side of intelligent algorithms and cognitive machines. *International Review of Administrative Sciences*, *90*(2), 353–368.

Corfield, G. (2023, September 14). 'British judge uses "jolly useful" ChatGPT to write ruling.' *The Telegraph*. https://www.telegraph.co.uk/business/2023/09/14/british-judge-uses-jolly-useful-chatgpt-to-write-ruling

Courts and Tribunals Judiciary. (2023). *Artificial intelligence (AI): Judicial guidance*. https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf

Creemers, R., & Webster, G. (2021). *Translation: Personal information protection law of the People's Republic of China—Effective Nov. 1, 2021*. DigiChina. https://digichina.stanford.edu/work/translation-personal-information-protection-law-of-the-peoples-republic-of-china-effective-nov-1-2021

Cyberspace Administration of China. (2021). *Internet information service algorithmic recommendation management provisions*.

Cyberspace Administration of China. (2023). *Interim measures for the management of generative artificial intelligence services*.

Department for Science, Innovation and Technology. (2023). *A pro-innovation approach to AI regulation: Government response*. UK Government. https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response

Department for Science, Innovation and Technology. (2024). *New records detail how AI helps government make quick, accurate decisions to boost trade, speed up responses and more.* UK Government. https://www.gov.uk/government/news/new-records-detail-how-ai-helps-government-make-quick-accurate-decisions-to-boost-trade-speed-up-responses-and-more

Department for Science, Innovation and Technology. (2025a). *AI opportunities action plan* (CP 1241). UK Government. https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan

Department for Science, Innovation and Technology. (2025b). *State of digital government review* (CP 1251). UK Government. https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review

Department for Science, Innovation and Technology. (2025c). *A blueprint for modern digital government* (CP 1252). UK Government. https://www.gov.uk/government/publications/a-blueprint-for-modern-digital-government

Department for Science, Innovation and Technology. (2025d). *Prime minister sets out blueprint to turbocharge AI*. UK Government. https://www.gov.uk/government/news/prime-minister-sets-out-blueprint-to-turbocharge-ai

Department for Science, Innovation and Technology & AI Safety Institute. (2025). *International AI safety report 2025* (DSIT research paper series no. 2025/001). Crown Copyright. https://assets.publishing.service.gov.uk/media/679a0c48a77d250007d313ee/International_AI_Safety_Report_2025_accessible_f.pdf

Department for Work and Pensions. (2024). *DWP annual report and accounts 2023 to 2024*. UK Government. https://www.gov.uk/government/publications/dwp-annual-report-and-accounts-2023-to-2024

Department of Education. (2023). *The impact of AI on UK jobs and training*. https://assets.publishing.service.gov.uk/media/656856b8cc1ec500138eef49/Gov.UK_Impact_of_AI_on_UK_Jobs_and_Training.pdf

Department of Homeland Security. (2024). *AI at DHS: A deep dive into our use case inventory*. UK Government. https://www.dhs.gov/archive/news/2024/12/16/ai-dhs-deep-dive-our-use-case-inventory

Digital Watch Observatory. (2023). *Indian judge used ChatGPT in a criminal case*. https://dig.watch/updates/indian-judge-used-chatgpt-in-a-criminal-case

Edwards, L., & Veale, M. (2018). *Enslaving the algorithm: From a "right to an explanation" to a "right to better decisions"?* arXiv. https://arxiv.org/abs/1803.07540

Engel, C., Linhardt, L., & Schubert, M. (2024). Code is law: How COMPAS affects the way the judiciary handles the risk of recidivism. *Artificial Intelligence and Law, 33*, 383–404. https://doi.org/10.1007/s10506-024-09389-8

Espinoza, J., & Dubois, L. (2025, March 25). EU lawmakers warn against 'dangerous' moves to water down AI rules. *Financial Times*. https://www.ft.com/content/9051af42-ce3f-4de1-9e68-4e0c1d1de5b5

European Commission. (2000). *Communication from the Commission on the precautionary principle* (COM(2000) 1 final). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0001:FIN:EN:PDF

European Commission. (2019). *Ethics guidelines for trustworthy AI*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission. (2020a). *White paper on artificial intelligence: A European approach to excellence and trust*. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

European Commission. (2020b). *Sectoral considerations on the policy and investment recommendations for*

trustworthy artificial intelligence. Publications Office of the European Union. https://futurium.ec.europa.eu/en/european-ai-alliance/community-content/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai

European Commission. (2020c). *Digitalisation of justice in the European Union: A toolbox of opportunities* (COM(2020) 710 final). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2020:710:FIN

European Commission. (2021). *Coordinated plan on artificial intelligence 2021 review*. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

European Commission. (2023). *AI watch: European landscape on the use of Artificial intelligence by the public sector*. https://interoperable-europe.ec.europa.eu/sites/default/files/custom-page/attachment/2023-02/JRC129301_01_AI_watch.pdf

European Commission. (2024a). *EU study calls for strategic AI adoption to transform public sector services*. https://digital-strategy.ec.europa.eu/en/library/eu-study-calls-strategic-ai-adoption-transform-public-sector-services

European Commission. (2024b). *The potential of generative AI for the public sector: Current use, key questions and policy considerations*. https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/document/potential-generative-ai-public-sector-current-use-key-questions-and-policy-considerations

European Commission. (2025, February 11). *EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence* [Press release]. https://digital-strategy.ec.europa.eu/en/news/eu-launches-investai-initiative-mobilise-eu200-billion-investment-artificial-intelligence

Foucault, M. (1985). *Discourse and truth: The problematization of parrhesia*. Northwestern University Press.

Foy, H., & Moens, B. (2025, February 14). EU scales back tech rules to boost AI investment, says digital chief. *Financial Times*. https://www.ft.com/content/fde53886-4295-4066-a704-b8cf5f388800

Future of Life Institute. (2023). *Pause giant AI experiments: An open letter*. https://futureoflife.org/open-letter/pause-giant-ai-experiments

Gallo, V., & Nair, S. (2023). *The UK's framework for AI regulation*. Deloitte. https://www2.deloitte.com/uk/en/blog/emea-centre-for-regulatory-strategy/2024/the-uks-framework-for-ai-regulation.html

Gong, J., & Dorwart, H. (2024). *AI governance in China: Strategies, initiatives, and key considerations*. Bird & Bird. https://www.twobirds.com/en/insights/2024/china/ai-governance-in-china-strategies-initiatives-and-key-considerations

Greenacre, M. (2024, December 5). EU is 'losing the narrative battle' over AI Act, says UN adviser. *Science|Business*. https://sciencebusiness.net/news/ai/eu-losing-narrative-battle-over-ai-act-says-un-adviser

Gutiérrez, J. (2024). A critical appraisal of large language models in judicial decision-making. In R. Paul, E. Carmel, & J. Cobbe (Eds.), *Handbook on public policy and artificial intelligence* (p. 328–338) Edward Elgar.

Gutiérrez, J. D., & Muñoz-Cadena, S. (2024). *Algorithmic transparency in the public sector: A state-of-the-art report of algorithmic transparency instruments*. Global Partnership on Artificial Intelligence. https://wp.oecd.ai/app/uploads/2024/12/14-Algorithmic-Transparency-in-the-Public-Sector-A-state-of-the-art-report-of-algorithmic-transparency-instruments.pdf

Hacker, P. (2023). The European AI liability directives—Critique of a half-hearted approach and lessons for the future. *Computer Law & Security Review, 51*, Article 105871.

Hadwick, D., & Lan, S. (2021). Lessons to be learned from the Dutch childcare allowance scandal: A comparative review of algorithmic governance by tax administrations in the Netherlands, France and Germany. *World Tax Journal, 13*(4), 609–645. https://www.ibfd.org/shop/journal/lessons-be-learned-dutch-childcare-allowance-scandal-comparative-review-algorithmic

Heikkilä, M. (2022, March 29). Dutch scandal serves as a warning for Europe over risks of using algorithms. *Politico*. https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms

Hobbs, C. (2020, April 8). The EU as a digital regulatory superpower: Implications for the United States. *European Council on Foreign Relations*. https://ecfr.eu/article/commentary_the_eu_as_a_digital_regulatory_superpower_implications_for_the_u

House of Lords. (2024). *Public authority algorithmic and automated decision-making systems bill*. https://bills.parliament.uk/bills/3760

Howell, B. (2023). *The precautionary principle, safety regulation, and AI: This time, it really is different*. American Enterprise Institute. https://www.aei.org/research-products/report/the-precautionary-principle-safety-regulation-and-ai-this-time-it-really-is-different

Hsu, S. (2021, November 1). China's regulatory clampdown on Big Tech: Motivations and the American response. *Institute for China-America Studies*. https://chinaus-icas.org/research/chinas-regulatory-clampdown-on-big-tech

Institute for the Future of Work. (2025). *Final report of the Pissarides review into the future of work and wellbeing*. https://cdn.prod.website-files.com/64d5f73a7fc5e8a240310c4d/679baff5f45270c64b9bda11_TPR-FinalReport-26-01-25v4.pdf

Jung, C., & Desikan, B. S. (2024). *Transformed by AI: How generative artificial intelligence could affect work in the UK—And how to manage it*. IPPR. https://www.ippr.org/articles/transformed-by-ai

Kaminski, M. E. (2018). *The right to explanation, explained*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3196985

Katsh, E., & Rabinovich-Einy, O. (2017). *Digital justice: Technology and the internet of disputes*. Oxford University Press.

Kippin, S., & Cairney, P. (2022). The Covid-19 exams fiasco across the UK: Four nations and two windows of opportunity. *British Politics*, *17*, 1–23. https://doi.org/10.1057/s41293-021-00162-y

Kretschmer, M., Kretschmer, T., Peukert, A., & Peukert, C. (2023). *The risks of risk-based AI regulation: Taking liability seriously*. arXiv. https://arxiv.org/abs/2311.14684

Liao, K. (2025). *Uphold the predominant position of political security*. (Original work published April 2022). https://www.strategictranslation.org/articles/chapter-six-persevere-in-placing-political-security-in-the-predominant-position

Loricchio, L., & Wallace, C. (2024, October 21). Transparency, oversight urged for IRS artificial intelligence. *Tax Notes*. https://www.taxnotes.com/featured-news/transparency-oversight-urged-irs-artificial-intelligence/2024/10/21/7m6nv

Lovely, G. (2024, September 11). SAG-AFTRA and women's groups urge Gavin Newsom to sign AI safety bill. *The Verge*. https://www.theverge.com/2024/9/11/24242142/sagaftra-ai-now-gavin-newsom-safety-sb-1047-letters

Majone, G. (1994). The rise of the regulatory state in Europe. *West European Politics*, *17*(3), 77–101.

Malloy, D. (2023, June 15). The world's regulatory superpower is taking on a regulatory nightmare: Artificial intelligence. *Atlantic Council*. https://www.atlanticcouncil.org/blogs/new-atlanticist/the-worlds-regulatory-superpower-is-taking-on-a-regulatory-nightmare-artificial-intelligence

Mark, D., McInerney, T., & Morison, J. (2024). Regulating automated decision-making in the justice system: What is the problem? In R. Paul, E. Carmel, & J. Cobbe (Eds.), *Handbook on public policy and artificial intelligence*. Edward Elgar Publishing. https://www.elgaronline.com/edcollchap/book/9781803922171/book-part-9781803922171-34.xml

McKinsey & Company. (2022). *The state of AI in 2022—And a half-decade in review.* https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

McKinsey & Company. (2025). *The state of AI: How organizations are rewiring to capture value.* https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Mergel, I., Dickinson, H., Stenvall, J., & Gasco, M. (2023). Implementing AI in the public sector. *Public Management Review.* Advance online publication. https://doi.org/10.1080/14719037.2023.2231950

Miller, P., & Rose, N. (2008). *Governing the present: Administering economic, social and personal life.* Polity Press.

Milmo, D. (2025, February 3). AI systems could be 'caused to suffer' if consciousness achieved, says research. *The Guardian.* https://www.theguardian.com/technology/2025/feb/03/ai-systems-could-be-caused-to-suffer-if-consciousness-achieved-says-research

Ministry of Justice. (2025, July 31). *AI action plan for justice.* GOV.UK. https://www.gov.uk/government/publications/ai-action-plan-for-justice/ai-action-plan-for-justice

Morison, J., & Harkens, A. (2019). Re-engineering justice? Robot judges, computerized courts and (semi) automated legal decision-making. *Legal Studies, 39*(4), 618–635.

Morison, J., & McInerney, T. (2025). When should a computer decide? Judicial decision-making in the age of automation, algorithms and generative artificial intelligence. In S. Turenne & M. Moussa (Eds.), *Research handbook on judging and the judiciary.* Elgar; Routledge. https://www.researchgate.net/publication/378610409_When_should_a_computer_decide_Judicial_decision-making_in_the_age_of_automation_algorithms_and_generative_artificial_intelligence

National Audit Office. (2024). *Use of artificial intelligence in government.* https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1).* U.S. Department of Commerce.

National Technical Committee. (2024). *AI safety governance framework.* https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf

Nazzaro, M. (2024, September 9). AI employees voice support for California regulation bill. *The Hill.* https://thehill.com/policy/technology/4869225-ai-employees-support-california-ai-bill

Neidig, H. (2018, January 10). Chamber of Commerce President warns against growing 'techlash.' *The Hill.* https://thehill.com/policy/technology/368331-chamber-of-commerce-president-warns-against-growing-techlash

New York State Legislature. (2024). *An act to amend the executive law in relation to requiring state agencies to evaluate and report on their use of artificial intelligence software* (Assembly Bill A.5672).

Office of Management and Budget. (2020). *M-21-06: Guidance for regulation of artificial intelligence applications.* Executive Office of the President.

Office of Management and Budget. (2025). *Memorandum M-25-21: Advancing governance, innovation, and risk management for agency use of artificial intelligence.* Executive Office of the President of the United States.

Office of Science and Technology Policy. (2020). *American artificial intelligence initiative: Year one annual report.* Executive Office of the President.

Office of Science and Technology Policy. (2022). *Blueprint for an AI Bill of Rights: Making automated systems work for the American people.* The White House. https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights

OpenAI. (2024). *Learning to reason with LLMs.* https://openai.com/index/learning-to-reason-with-llms

OpenAI. (2025). *Announcing the stargate project.* https://openai.com/index/announcing-the-stargate-project

Oremus, W. (2023, April 4). The AI backlash is here. It's focused on the wrong things. *The Washington Post.* https://www.washingtonpost.com/technology/2023/04/04/musk-ai-letter-pause-robots-jobs

Organisation for Economic Co-operation and Development. (2023). *OECD digital government index*. https://www.oecd.org/en/publications/2023-oecd-digital-government-index_1a89ed5e-en.html

Organisation for Economic Co-operation and Development. (2024). *Governing with artificial intelligence: Are governments ready?* (OECD Artificial Intelligence Papers, No. 20). https://doi.org/10.1787/26324bc2-en

Organisation for Economic Co-operation and Development, & UNESCO. (2024). *G7 toolkit for artificial intelligence in the public sector*. https://doi.org/10.1787/421c1244-en

Oxford Insights. (2024). *Government AI readiness index 2024*. https://oxfordinsights.com/ai-readiness/ai-readiness-index

Perrigo, B. (2023, August 1). E.U.'s AI regulation could be softened after pushback from biggest members. *TIME*. https://time.com/6338602/eu-ai-regulation-foundation-models

Pew Research Center. (2023). *What the data says about Americans' views of artificial intelligence*. https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence

Pinchai, S., Hassabis, D., & Kavukcuoglu, K. (2025). *Introducing Gemini: Redefining AI with integrated reasoning*. Google. https://blog.google/technology/google-deepmind/google-gemini-ai-update-december-2024

Public Law Project. (2023). *Written evidence to the Justice and Home Affairs Committee (NTL0046): Technology rules? The advent of new technologies in the justice system*. https://committees.parliament.uk/writtenevidence/39761/html

PwC. (2025). *AI adoption could boost global GDP by an additional 15 percentage points by 2035, as global economy is reshaped*. https://www.pwc.com/gx/en/news-room/press-releases/2025/ai-adoption-could-boost-global-gdp-by-an-additional-15-percentage.html

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). (2016). *Official Journal of the European Union*, L 119.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (AI Act). (2024). *Official Journal of the European Union*, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Rekenkamer, A. (2024). Netherlands Court of audit, *focus op AI bij de rijksoverheid*. https://www.rekenkamer.nl/publicaties/rapporten/2024/10/16/focus-op-ai-bij-de-rijksoverheid

Roberts, H., Cowls, J., Morley, J., Taddeo, M., & Floridi, L. (2021). The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation. In L. Floridi (Ed.), *Ethics, governance, and policies in artificial intelligence* (pp. 23–51). Springer.

Roose, K. (2023, May 30). A.I. poses 'risk of extinction,' industry leaders warn. *The New York Times*. https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html

Ross v. United States, 23-CM-1067 (D.C. Ct. App) (2025). https://law.justia.com/cases/district-of-columbia/court-of-appeals/2025/23-cm-1067.html

Schneider, I. (2025). *Reclaiming digital sovereignty: The EU's role in the geopolitics of digital governance* (Policy Paper No. 1). Center for the Governance of Change. https://static.ie.edu/CGC/CGC_ReclaimingDigitalSovereignty_PolicyPaper.pdf

Sénécat, A. (2023, December 5). The use of opaque algorithms facilitates abuses within public services. *Le Monde*. https://www.lemonde.fr/en/les-decodeurs/article/2023/12/05/the-use-of-opaque-algorithms-facilitates-abuses-within-public-services_6314051_8.html

Sheehan, M. (2023, July 12). China's AI regulations and how they get made. *Carnegie Endowment for International Peace*. https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made?lang=en

Singer, J., & Sheehan, M. (2025). *China's AI policy at the crossroads: Balancing development and control in the DeepSeek era*. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2025/07/chinas-ai-policy-in-the-deepseek-era?lang=en

Sourdin, T. (2021). *Judges, technology and artificial intelligence: The artificial judge*. Edward Elgar.

Stacey, K. (2023, October 23). UK officials use AI to decide on issues from benefits to marriage licences. *The Guardian*. https://www.theguardian.com/technology/2023/oct/23/uk-officials-use-ai-to-decide-on-issues-from-benefits-to-marriage-licences

Stanford Institute for Human-Centered Artificial Intelligence. (2023). *AI index report 2023*. https://hai.stanford.edu/ai-index/2023-ai-index-report

Stanford Institute for Human-Centered Artificial Intelligence. (2025). *AI index report 2025*. https://hai.stanford.edu/ai-index/2025-ai-index-report

Sun, Y., & Zeng, J. (2024, April 22). China's interim measures for the management of generative AI services: A comparison between the final and draft versions of the text. *Future of Privacy Forum*. https://fpf.org/blog/chinas-interim-measures-for-the-management-of-generative-ai-services-a-comparison-between-the-final-and-draft-versions-of-the-text

Supreme People's Court. (2022). *Chinese courts must implement AI system by 2025*. https://english.court.gov.cn/2022-12/12/c_838810.htm

Taylor, L. (2023, February 3). Colombian judge says he used ChatGPT in ruling. *The Guardian*. https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling

Thierer, A. D. (2016). *Permissionless innovation: The continuing case for comprehensive technological freedom*. Mercatus Center at George Mason University.

Trump, D. J. (2019). *Maintaining American leadership in artificial intelligence* (Executive Order No. 13859). Federal Register. https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence

Trump, D. J. (2025). *Remarks on artificial intelligence infrastructure development and an exchange with reporters*. The American Presidency Project. https://www.presidency.ucsb.edu/documents/remarks-artificial-intelligence-infrastructure-development-and-exchange-with-reporters

UNESCO. (2024). *UNESCO global judges' initiative: Survey on the uses of AI systems by judicial operators*. https://unesdoc.unesco.org/ark:/48223/pf0000389786

United Nations. (2016). *UN e-government survey 2016*. https://publicadministration.un.org/egovkb/en-us/reports/un-e-government-survey-2016

US Department of Defense. (2018). *AI, China, Russia, and the global order: Technological, political, global, and creative perspectives*. Defense Technical Information Center. https://apps.dtic.mil/sti/citations/AD1066673

US House of Representatives. (2025). *Majority memorandum for May 13, 2025, Committee on Energy and Commerce markup*. https://docs.house.gov/meetings/IF/IF00/20250513/118261/HMKP-119-IF00-20250513-SD003.pdf

Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, *19*, 255–274.

Viljoen, S. (2021). The promise and limits of lawfulness: Inequality, law, and the techlash. *Journal of Social Computing*, *2*(3), 284–296.

Virginia General Assembly. (2025). *High-risk artificial intelligence developer and deployer act* (House Bill 2094).

Vogel, D. (2012). *The politics of precaution: Regulating health, safety, and environmental risks in Europe and the United States*. Princeton University Press.

Von der Leyen, U. (2025, May 20). *Speech by President von der Leyen at the Annual EU Budget Conference 2025* [Speech transcript]. European Commission. https://ec.europa.eu/commission/presscorner/detail/nl/speech_25_1284

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.

Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence, 10*(2), 1–37. https://doi.org/10.2478/jagi-2019-0002

Webster, G., Creemers, R., Kania, E., & Triolo, P. (2017). *Full translation: China's "New Generation Artificial Intelligence Development Plan" (2017)*. DigiChina. https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017

Weiss-Blatt, N., Thierer, A., & Barkley, T. (2024). *The AI technopanic and its effects: A primer*. Abundance Institute. https://api.slash.am/storage/v1/object/public/page_uploads/ed9aed01-09a1-4d95-be93-405ccebe50dd/articles/1719857803918/the-ai-technopanic-and-its-effects.pdf

White House. (1997). *A framework for global electronic commerce*. https://clintonwhitehouse4.archives.gov/WH/New/Commerce

White House. (2025a). *Fact sheet: President Donald J. Trump takes action to enhance America's AI leadership*. https://www.whitehouse.gov/fact-sheets/2025/01/fact-sheet-president-donald-j-trump-takes-action-to-enhance-americas-ai-leadership

White House. (2025b). *White House releases new policies on federal agency AI use and procurement*. https://www.whitehouse.gov/articles/2025/04/white-house-releases-new-policies-on-federal-agency-ai-use-and-procurement

White House. (2025c. *Winning the race: America's AI Action Plan*. https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan

White House. (2025d). *Executive order 14179: Removing barriers to American leadership in artificial intelligence*. Federal Register. https://www.federalregister.gov/documents/2025/01/31/2025-02172/removing-barriers-to-american-leadership-in-artificial-intelligence

White House. (2025e). *Winning the race: America's AI Action Plan*. https://www.whitehouse.gov/ai-action-plan

Williams, M., & Sibley, L. (2022, May 20). *AI report 2022*. Marks & Clerk. https://www.marks-clerk.com/insights/news/102jvti-ai-report-2022

xAI. (2025). *The age of reasoning agents*. https://x.ai/blog/grok-3

Xia, Y. (2024). Research on judicial trial practice issues of internet courts in China. *Science of Law Journal, 3*, 174–179.

Ye, J. (2023, July 13). China says generative AI rules to apply only to products for the public. *Reuters*. https://www.reuters.com/technology/china-issues-temporary-rules-generative-ai-services-2023-07-13

Yeung, K. (2017). Algorithmic regulation: A critical interrogation. *Regulation & Governance, 12*(4), 505–523. https://doi.org/10.1111/rego.12158

Zeng, J. (2020). Artificial intelligence and China's authoritarian governance. *International Affairs, 96*(6), 1441–1458.

Zuiderveen Borgesius, F., & van Bekkum, M. (2021, September 23). Digital welfare fraud detection and the Dutch SyRI judgment. *International Association of Privacy Professionals*. https://iapp.org/news/a/digital-welfare-fraud-detection-and-the-dutch-syri-judgment

About the Authors

David Mark has completed his PhD within the Leverhulme Interdisciplinary Network on Algorithmic Solutions (LINAS) programme at Queen's University Belfast. His work explores the intricacies of safety, security, and legal compliance in AI decision-making systems. A former barrister, he has an MSc in software development and a wider interest in the transformative effects of algorithmic technologies on the legal sector.

John Morison is a professor of jurisprudence in the School of Law, Queen's University Belfast, and a member of the Royal Irish Academy. He has published widely in constitutional law and theory and on the impact of new technology. Currently, he runs the Leverhulme Interdisciplinary Network on Algorithmic Systems (LINAS), which funds 30 PhD researchers in an interdisciplinary research programme.