America's AI action plan and global governance: convergences, divergences, and theoretical implications

Darci de Borba Santos Júnior, Rafael Alfonso Brinkhues

https://doi.org/10.1590/0034-761220250584x

Submitted on: 2026-04-27
Posted on: 2026-04-28 (version 1)
(YYYY-MM-DD)

FGV EBAPE

RAP
REVISTA DE ADMINISTRAÇÃO PÚBLICA

Article

America's AI action plan and global governance: convergences, divergences, and theoretical implications

Darci de Borba Santos Júnior

Instituto de Pesquisa Econômica Aplicada, Diretoria de Desenvolvimento Institucional, Brasília, DF, Brazil. ORCID: https://orcid.org/0000-0001-9851-5034

Rafael Alfonso Brinkhues

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS), Campus Viamão, RS, Brazil. ORCID: https://orcid.org/0000-0002-9367-5829

Presidency of the Federative Republic of Brazil, Brasília, DF, Brazil

Abstract

This study critically evaluates America's AI Action Plan (AAP) against best practices in artificial intelligence regulation and governance identified in recent scientific literature. Based on an integrative review of 100 peer-reviewed articles indexed in Scopus, the analysis contrasts the AAP's three pillars, innovation, infrastructure, and diplomacy, with key governance dimensions, namely accountability, sustainability, and multilateralism. The findings indicate selective convergences, especially in cybersecurity, technological sovereignty, and standardization, alongside structural divergences concerning rights protection, environmental responsibility, and global cooperation. Drawing on these

DOI: https://doi.org/10.1590/0034-761220250584x
Submitted on October 29, 2025, and accepted for publication on March 07, 2026.
[Original version]

results, the paper proposes a theoretical framework that conceptualizes AI governance as a multidimensional and adaptive regime integrating innovation, regulation, ethics, sustainability, and epistemic co-production. The study offers practical value by informing policymakers and regulators on how to strengthen adaptive and risk-based regulatory design, and it contributes to the academic debate by framing AI governance as an emergent global normative system. The recommendations are intentionally delimited to public policy and governance arrangements, with implications for global coordination and the design of national AI strategies, rather than prescribing technical solutions for specific models or systems. Limitations include the exclusive use of Scopus and the focus on a single national strategy; future research should expand comparative analyses across regions and examine longitudinal transformations in AI governance regimes.

Keywords: artificial intelligence governance, regulation, global policy, sustainability, international cooperation.

Plano de ação de IA dos Estados Unidos e governança global: convergências, divergências e implicações teóricas

Resumo

Este estudo avalia criticamente o America's AI Action Plan (AAP) à luz de boas práticas em regulação e governança de inteligência artificial identificadas na literatura científica recente. Com base em uma revisão integrativa de 100 artigos revisados por pares e indexados na Scopus, a análise contrasta os três pilares do AAP – inovação, infraestrutura e diplomacia – com dimensões centrais de governança, especificamente accountability, sustentabilidade e multilateralismo. Os resultados indicam convergências seletivas, com maior aderência em temas como cibersegurança, soberania tecnológica e padronização, ao passo que evidenciam divergências estruturais no que se refere à proteção de direitos, à responsabilização por impactos socioambientais e ao fortalecimento de mecanismos de cooperação internacional. A partir desses achados, o artigo propõe um framework teórico que conceitua a governança de IA como um regime multidimensional e adaptativo, integrando inovação, regulação, ética, sustentabilidade e coprodução epistêmica. O estudo oferece valor prático ao subsidiar formuladores de políticas públicas e reguladores na qualificação de desenhos regulatórios adaptativos e orientados a risco, e contribui para o debate acadêmico ao enquadrar a governança de IA como um sistema normativo global emergente. As recomendações são deliberadamente delimitadas a arranjos de política

pública e governança, com implicações para a coordenação global e para o desenho de estratégias nacionais de IA, sem prescrever soluções técnicas para modelos ou sistemas específicos. Entre as limitações, destacam-se o uso exclusivo da Scopus e o foco em uma única estratégia nacional; pesquisas futuras devem ampliar comparações inter-regionais e examinar transformações longitudinais em regimes de governança de IA.

Palavras-chave: governança da inteligência artificial, regulação, política global, sustentabilidade, cooperação internacional.

Plan de acción de IA de Estados Unidos y gobernanza global: convergencias, divergencias e implicaciones teóricas

Resumen

Este estudio evalúa críticamente el America's AI Action Plan (AAP) a la luz de las mejores prácticas de regulación y gobernanza de la inteligencia artificial identificadas en la literatura científica reciente. A partir de una revisión integradora de 100 artículos revisados por pares e indexados en Scopus, el análisis contrasta los tres pilares del AAP – innovación, infraestructura y diplomacia – con dimensiones clave de la gobernanza, en particular la rendición de cuentas, la sostenibilidad y el multilateralismo. Los hallazgos muestran convergencias selectivas, con mayor alineación en ámbitos como la ciberseguridad, la soberanía tecnológica y la estandarización, junto con divergencias estructurales en lo relativo a la protección de derechos, la responsabilidad por impactos socioambientales y el fortalecimiento de mecanismos de cooperación internacional. Con base en estos resultados, el artículo propone un marco teórico que conceptualiza la gobernanza de la IA como un régimen multidimensional y adaptativo que integra innovación, regulación, ética, sostenibilidad y coproducción epistémica. El estudio aporta valor práctico al orientar a responsables de políticas públicas y reguladores sobre cómo fortalecer diseños regulatorios adaptativos y basados en riesgo, y contribuye al debate académico al enmarcar la gobernanza de la IA como un sistema normativo global emergente. Las recomendaciones se delimitan intencionalmente a políticas públicas y arreglos de gobernanza, con implicaciones para la coordinación global y el diseño de estrategias nacionales de IA, sin prescribir soluciones técnicas para modelos o sistemas específicos. Entre las limitaciones, se destacan el uso exclusivo de Scopus y el enfoque en una única estrategia nacional; investigaciones futuras deberían ampliar los análisis

comparativos entre regiones y examinar transformaciones longitudinales en los regímenes de gobernanza de la IA.

Palabras clave: gobernanza de la inteligencia artificial, regulación, política global, sostenibilidad, cooperación internacional.

1. INTRODUCTION

The recent acceleration of artificial intelligence (AI) capabilities has repositioned technological governance at the center of the international political agenda. Observers highlight that the race for standards and transparency is likely to shape global norms, given the transnational nature of value chains and the broad impacts of AI (Lebeaux, 2025). In this context, AI regulation becomes unavoidable to mitigate systemic risks, provide legal predictability, and sustain public trust and competitiveness (Bannister & Connolly, 2020; Kaminski, 2023; Koshiyama et al., 2024). Coherent national strategies coordinate actors and prevent both regulatory vacuums and ethics washing, aligning innovation with democratic values and fundamental rights (Papyshev & Yarime, 2023; Radu, 2021).

Against this backdrop, (The White House, 2025) released the America's AI Action Plan (AAP), a governmental roadmap that structures federal action around three pillars: (i) accelerating innovation; (ii) building AI infrastructure; and (iii) leading in international diplomacy and security. The document is premised on geopolitical competition for technological dominance and sets the explicit goal of sustaining U.S. leadership through incentives for innovation, regulatory streamlining, and the international projection of the American technological stack (models, hardware, and standards). AI governance refers to the framework of principles, processes, and institutions guiding the development, evaluation, and use of systems in both the public and private sectors (Bannister & Connolly, 2020; Wirtz et al., 2020).

Although the AAP signals investments in R&D, infrastructure, and security, it also articulates a regulatory inflection by reducing red tape and revisiting provisions considered "onerous" for innovation, including guidance to re-evaluate secondary legislation and the regulatory framework itself (The White House, 2025). In parallel, it proposes revising the NIST AI RMF, removing references to misinformation, diversity, equity and inclusion (DEI), and climate issues, while emphasizing the protection of

freedom of expression in frontier models (The White House, 2025). These directions stand in tension with the literature, which stresses: (i) the centrality of accountability, risk assessment, and audits as counterbalances to innovation (Bell et al., 2023; Koshiyama et al., 2024; Wirtz et al., 2020); (ii) the limits of requiring "human oversight" without adequate institutional design (Green, 2022); and (iii) the importance of international regulatory convergence (Kaminski, 2023; Wörsdörfer, 2024). In sum, the research problem emerges as how to balance accelerated promotion of innovation with risk mitigation in the design proposed by the AAP.

In this study, AI governance is defined as the set of norms and institutions that shape how AI is built and deployed, as well as the associated policy and research efforts intended to steer these trajectories toward socially desirable outcomes (Dafoe, 2023). Under this operationalization, "governance" is treated as an institutional and policy arrangement that allocates authority, sets accountability expectations, and coordinates public and private actors across the AI lifecycle, rather than as a purely technical debate about model architectures or performance (Marchiori et al., 2025). Accordingly, the analytical scope is intentionally bounded to public policy and governance arrangements expressed in national strategies and their implementation logic, and it is examined through a closed set of dimensions that recur in the contemporary literature on AI governance and global governance regimes: regulatory oversight and accountability, including transparency, responsibility, and rights protection (Jobin et al., 2019); infrastructure, standard setting, and technological sovereignty as determinants of state capacity and control over AI ecosystems (Veale et al., 2023); ethical and social legitimacy as a condition for governance effectiveness (Jobin et al., 2019); environmental sustainability as a policy constraint for scaling AI infrastructures (Marchiori et al., 2025); and multilateralism as the coordination challenge of aligning domestic strategies with international regimes, standards, and cooperative mechanisms (Roberts et al., 2024). This set of dimensions is adopted to ensure conceptual precision and comparability, enabling a structured assessment of convergences, divergences, and tradeoffs across the plan's pillars while maintaining alignment with the theoretical framework advanced in the paper (Dafoe, 2023; Veale et al., 2023)

Comparative literature shows that national AI strategies vary in terms of the role of the state, the degree of prescriptiveness, and the prioritization of security and rights (Albous et al., 2025; Djeffal et al., 2022; Papyshev & Yarime, 2023; Radu, 2021). The

AAP may serve as a trajectory marker in the global AI regime by prioritizing pro-growth deregulation, energy and computational infrastructure, and the international projection of standards and export controls. The document also makes explicit its intention to diffuse the American technological stack among allies, counterbalance Chinese influence in governance forums, and reinforce export control mechanisms and chip localization verification (The White House, 2025). Considering that standards and norms often spread through market power and technical diplomacy (Ingersleben-Seip, 2023), critically assessing the AAP is essential for understanding regulatory convergences and divergences and their effects on innovation, competition, and safeguards.

Accordingly, this study asks: **How can the America's AI Action Plan (AAP) be evaluated considering the best practices of AI regulation and governance identified in the scientific literature?** The objective is to critically analyze the AAP, identifying its consistencies, gaps, and trade-offs across the three pillars of innovation, infrastructure, and diplomacy, and to assess its regulatory implications for the global AI ecosystem, particularly regarding standards, audits, and risk-based governance, through a systematic review of the literature on AI regulation and governance (Snyder, 2019).

This article offers three contributions to the field. First, it develops an analytical framework that integrates multiple theoretical dimensions of AI governance, enabling a structured assessment of regulatory coherence across national and global contexts. Second, it proposes an adherence and alignment model that allows the quantification of empirical and normative consistency among AI governance approaches, facilitating comparative evaluation of policy designs. Third, it advances a theoretical abstraction of AI governance as a multidimensional regime, highlighting how innovation, sovereignty, ethics, and cooperation interact to shape legitimacy and stability in global AI governance. Together, these contributions strengthen the theoretical and practical understanding of how regulatory models can reconcile technological dynamism with social safeguards, transparency, and international cooperation.

2. THEORETICAL FRAMEWORK

This study organizes the referential around four central constructs that recur in recent debates on AI regulation and governance and that are consistent with the analytical aim of evaluating a national strategy against best practices identified in literature. First,

legitimacy is treated as the condition under which AI governance arrangements are perceived as trustworthy and acceptable, particularly when they embed accountability, transparency, and safeguards that sustain public confidence in policy choices and public sector uses of AI (Bannister & Connolly, 2020; Wirtz et al., 2020; Yoganathan et al., 2025). Second, technological sovereignty is conceptualized as state capacity to secure strategic autonomy in AI ecosystems through control over critical infrastructures, supply chains, standards, and risk management capabilities, especially in contexts where AI is intertwined with national security and industrial strategy (Papyshev & Yarime, 2023; Russell & McGravey, 2025). Third, sustainability captures the incorporation of environmental externalities and long-run ecological constraints into AI governance, including the energy and carbon implications of large-scale compute, data centers, and digital infrastructure growth (Han et al., 2025; Koenig et al., 2023). Fourth, multilateralism refers to cooperative and polycentric coordination among states and other actors to reduce fragmentation, promote convergence of standards, and govern cross-border risks through shared norms and institutional arrangements (Ingersleben-Seip, 2023; Radu, 2021; Roberts et al., 2023; Robinson, 2025).

These constructs are theoretically interdependent, which allows the analysis to move beyond descriptive comparison and examine trade-offs. A first expected relation concerns the tension between sovereignty and multilateralism: sovereignty-oriented strategies may strengthen autonomy and enforcement capacity, yet they can also incentivize unilateral diffusion of standards and selective cooperation, increasing the risk of regulatory fragmentation (Ingersleben-Seip, 2023; Radu, 2021). A second relation links legitimacy and regulatory calibration: governance approaches that privilege accelerated innovation without commensurate accountability mechanisms tend to face legitimacy constraints, including reduced trust and stronger demand for state oversight, particularly when risks are salient to citizens and stakeholders (Bannister & Connolly, 2020; Wirtz et al., 2020; Yoganathan et al., 2025). A third relation concerns sustainability and infrastructure expansion: infrastructure-centric strategies may strengthen competitiveness and sovereignty, but sustainability pressures can constrain scale-up and reshape policy choices, affecting long-run feasibility and international acceptance (Han et al., 2025; Koenig et al., 2023). In combination, these relations provide a coherent basis for evaluating policy designs by examining whether, and how, strategies reconcile innovation

and accountability, autonomy and cooperation, and infrastructure growth and environmental responsibility (Papyshev & Yarime, 2023; Roberts et al., 2023).

From an evaluative standpoint, the framework implies that policy assessment should focus on observable governance features that instantiate each construct. Legitimacy is assessed through the presence and robustness of accountability instruments, audits, transparency obligations, and rights protections (Bannister & Connolly, 2020; Wirtz et al., 2020). Technological sovereignty is assessed through measures addressing compute, semiconductors, cybersecurity, standards capacity, and resilience of strategic infrastructures (Papyshev & Yarime, 2023; Russell & McGravey, 2025). Sustainability is assessed through the extent to which environmental externalities are recognized and governed in infrastructure policies and related regulatory choices (Han et al., 2025; Koenig et al., 2023). Multilateralism is assessed through commitments to international coordination, standard convergence, and cooperative institutions, including approaches that reduce cross-jurisdictional inconsistency (Ingersleben-Seip, 2023; Radu, 2021; Robinson, 2025; Veale et al., 2023). This structure supports a systematic evaluation of convergences, divergences, and the trade-offs embedded in national AI strategies.

2.1 Overview of the America's AI Action Plan

The AAP constitutes the most recent and comprehensive initiative of the United States government as an explicit attempt to consolidate the country's global leadership in the field of AI (The White House, 2025). The document is premised on the notion that technological supremacy is a matter of national security and, therefore, must be achieved through policies that accelerate innovation, strengthen infrastructure, and ensure diplomatic and military prominence in the international AI ecosystem. Inspired by narratives such as the space race, the AAP frames technological competition in AI as a watershed moment capable of reshaping the global balance of power, directly impacting the economy, and influencing international regulatory standards.

The AAP is structured around three central pillars. Pillar I – Accelerating AI Innovation focuses on creating an environment conducive to private-sector-led innovation by removing regulatory barriers, stimulating open-source and open-weight models, protecting freedom of expression in AI systems, and promoting adoption in strategic sectors such as health, defense, and public administration. It also includes

measures aimed at workforce reskilling, fostering advanced manufacturing, and strengthening a national ecosystem for AI evaluation, in addition to supporting applied scientific research.

Pillar II – Building American AI Infrastructure establishes that AI supremacy requires massive investments in energy, semiconductors, and data center infrastructure. The document calls for the simplification of environmental licensing processes, the restoration of domestic chip manufacturing, the construction of high-security data centers for military purposes, and the expansion of the technical workforce needed to sustain such growth. This pillar also underscores the need to strengthen the cybersecurity of critical infrastructures, promote secure-by-design technologies, and develop robust federal capacities for incident response involving AI.

Finally, Pillar III – Leading in AI Diplomacy and International Security aims to globally project U.S. influence by promoting the export of the American technological stack (hardware, software, models, and standards), counterbalancing China's influence in multilateral organizations, and reinforcing semiconductor export controls. Among the strategic actions outlined are: aligning protective measures with allied countries, conducting national risk assessments associated with frontier models, and investing in biosecurity to mitigate emerging risks in biotechnology and the malicious use of AI.

By articulating these three pillars, the AAP outlines a model of technology policy grounded in deregulation, technological nationalism, and the international projection of standards. It is, therefore, a document that combines industrial policy, national defense, and technological diplomacy, positioning the United States at the center of the dispute over the future direction of global AI governance.

3. METHOD

An integrative review was conducted (Snyder, 2019), appropriate for synthesizing conceptual and empirical evidence of different methodological natures on AI regulation and governance in the public sector. The purpose of the review was to inform the critical analysis of the AAP by comparing its pillars with recent literature. The primary search was conducted in Scopus, using the following parameterized string (titles, abstracts, and keywords), with filters for document type and language:

- TITLE-ABS-KEY("Artificial Intelligence" AND regulation AND Government)
    - AND PUBYEAR > 2019 AND PUBYEAR < 2027
    - AND (LIMIT-TO(DOCTYPE, "ar"))
    - AND (LIMIT-TO(LANGUAGE, "English"))

Thus, only articles (ar) published in English between 2020 and 2026 were considered. The review was restricted to the 2020–2026 period because AI governance and regulation entered an accelerated phase after 2019, when the OECD principles helped consolidate a shared baseline that rapidly translated into national strategies and risk-based policy debates, and because 2020 marks a clear inflection point in both technology and governance agendas with the scaling of foundation models and large language models, which reshaped concerns about accountability, safety, and systemic impacts and stimulated the consolidation of contemporary risk frameworks and policy toolkits.

The search was performed on August 3, 2025. Results were exported and managed in a spreadsheet for screening and extraction. It is worth noting that, since this is an integrative review, future complementarity with additional databases and snowballing techniques is admissible; however, Scopus was the primary source for this study. The extraction resulted in 351 documents, which were compared based on average H Index and average citations, as shown in Table 1.

Table 1

Number of extracted articles

Quadrant	Average H Index	Average Cited by	Quantity
Q1	141.69	18.61	192
Q2	42.78	18.72	82
Q3	21.65	6.20	51
Q4	13.00	0.88	24
Total	**92.06**	**15.52**	**351**

Source: Elaborated by the authors.

The Scopus database was chosen as the source for this study due to its comprehensive coverage, multidisciplinary scope, and high reliability in indexing peer-reviewed journals in the fields of public policy, technology, and governance. Scopus offers consistent metadata, citation metrics, and impact indicators that ensure the quality and traceability of academic evidence.

In addition to the Scopus database integrative review that resulted in a final corpus of 100 peer-reviewed articles, a citation snowballing procedure was applied to strengthen recall and reduce the risk of omitting influential studies not retrieved by keyword queries. Specifically, backward snowballing was conducted by screening the reference lists of the included articles, and forward snowballing was conducted by tracking subsequently citing publications. At each iteration, records were de-duplicated and subjected to the same eligibility criteria used in the primary screening, with iterative rounds continuing until marginal gains became negligible, as recommended in methodological guidance on snowballing for evidence synthesis (Wohlin, 2014).

To prioritize articles of greater impact and relevance, the sample was filtered using average citations and H Index. Articles with an H Index equal to or greater than 92.06 and those with citation counts equal to or greater than 15.52 were retained. A total of 200 articles that did not meet either criterion were excluded, leaving 151 articles in the sample (Table 2).

Table 2

Number of articles after quantitative criteria

Criterion	Average H Index	Average Cited by	Quantity
Average Cited by > 15.52	41.14	52.63	57
Average H Index > 92.06	279.92	3.56	62
Meet both criteria	188.41	47.06	32
Total	**170.39**	**31.30**	**151**

Source: Elaborated by the authors.

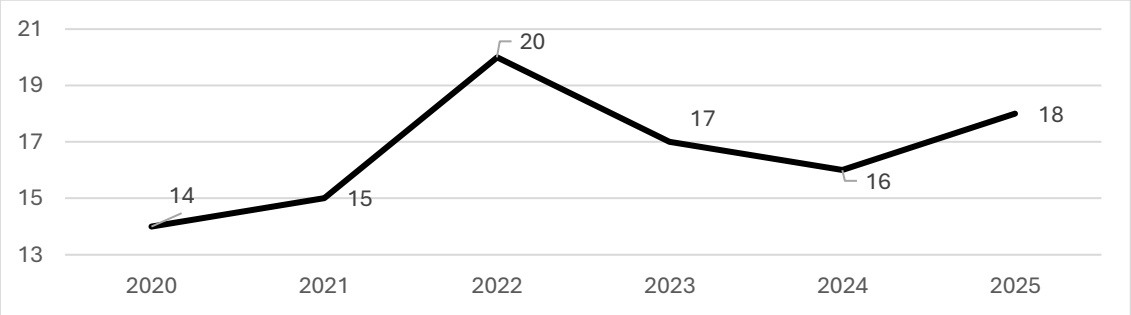
After reviewing the titles, abstracts, and keywords of the 151 remaining articles, 51 documents were excluded for not meeting the inclusion criteria: (1) addressing AI regulation or governance; (2) discussing national AI strategies or plans; and (3) analyzing infrastructure or economic impacts associated with AI policies. The following were also excluded: (1) purely technical studies; (2) sectoral analyses with no connection to governance or national policy; (3) perception or attitude surveys unrelated to regulatory and strategic debates; and (4) purely theoretical studies without application to policy, strategies, or infrastructure. As a result, 100 articles remained for analysis and review. The content analysis followed the stages proposed by (Bardin, 2008), which organizes the process into three phases: pre-analysis, material exploration, and treatment of results.

To ensure transparency and replicability, all bibliographic references used in this study were systematically managed through the Zotero platform. The complete reference collection, including all sources retrieved and reviewed during the integrative literature review, is publicly available (Zotero, 2025). This repository contains the database of articles analyzed in this research, serving as a complementary source for verification, citation tracing, and future updates related to artificial intelligence regulation and governance.

The Figure 1 shows the annual evolution of the number of publications analyzed between 2020 and 2025. A continuous increase is observed until 2022, when it peaks at 20 studies. From that point, there is a gradual decline in 2023 and 2024, followed by a slight recovery in 2025, indicating that academic interest in the topic remains high, albeit subject to fluctuations over time.

Figure 1

Number of articles over the years



Source: Elaborated by the authors.

In the pre-analysis phase, the selected articles were read and the units of analysis defined (title, abstract, results and conclusions). In the exploration phase, a system of deductive categories (Table 3) was constructed, aligned with the three pillars of the AAP (Innovation, Infrastructure, and Diplomacy/Security). Each article was coded according to the presence or absence of elements associated with these categories, and inductive subcategories were admitted when recurrent in the corpus. Finally, in the treatment of results, categories were quantified (frequency of occurrence) and qualitatively analyzed (critical interpretation), with triangulation between literature findings and the AAP's guidelines. This approach enables the assessment of convergences, divergences, and gaps between the U.S. policy and recent scientific production on AI regulation.

Table 3

Categories for content analysis

Pillar	Category	Description
Pillar I – Innovation and Regulation	C1 – Deregulation	References to regulatory flexibilization, removal of legal barriers, critique of rigid regulations
	C2 – Open models and standards	References to open-source, open-weight, audits, metrics, and technical standardization
	C3 – Protection of values and rights	Freedom of expression, ethics, diversity, inclusion, trust, and transparency
	C4 – Sectoral and public adoption	Initiatives to accelerate AI adoption in sectors such as health, government, and defense
	C5 – Labor and workforce training	Reskilling, education, and impacts of automation on workers
Pillar II – Infrastructure and Security	C6 – Technological infrastructure	Chips, data centers, energy, grids, and connectivity required for AI
	C7 – Cybersecurity and resilience	Protection of critical systems, incident response, secure-by-design
	C8 – Research and applied science	Databases, laboratories, scientific discoveries accelerated by AI
Pillar III – Diplomacy and Geopolitics	C9 – Export and global standards	Export of the American "stack" and international diffusion of standards
	C10 – Strategic competition	U.S.–China disputes and influence in multilateral forums
	C11 – Export controls	Sovereignty measures, semiconductors, global regulatory alignment
	C12 – Biosecurity and emerging risks	Mitigation of risks in biotechnology and malicious uses of AI

Source: Elaborated by the authors.

The coding system served as the analytical framework to organize the reading and interpretation of the articles included in the integrative review. Each publication was therefore examined according to the categories listed above, allowing the identification of trends, gaps, and implications for the debate on AI regulation in the context of the AAP.

3.1 Case selection

The AAP was selected as the exclusive unit of analysis for five complementary reasons. First, it is a highly recent federal strategy, released on July 23, 2025, which makes it a contemporaneous statement of regulatory priorities in a rapidly evolving governance landscape (The White House, 2025). Second, the AAP is issued by a jurisdiction with

pronounced geopolitical and market centrality in the AI ecosystem, which increases its agenda-setting capacity and the likelihood of demonstration effects across other national strategies and regulatory debates. Third, the document concentrates a high density of policy directives, explicitly listing more than 90 federal actions organized into three pillars, innovation, infrastructure, and international diplomacy and security, which supports a structured assessment of internal coherence and cross-dimensional trade-offs (The White House, 2025). Fourth, the AAP explicitly links domestic priorities to international diffusion by promoting exports of the U.S. AI technology stack and associated standards, which reinforces its potential influence on global rule-making and standard-setting processes. Fifth, this influence is consistent with existing U.S. governance infrastructures for technical standards and risk management, particularly NIST's coordinating role in AI standards and the international incorporation of the AI Risk Management Framework (Tabassi, 2023). To mitigate bias inherent in a single-case design, the limitations section acknowledges restricted generalizability, and the future research agenda advances systematic comparative extensions across regions and regulatory models, as well as longitudinal analyses of evolving AI governance regimes.

4. DATA ANALYSIS

A total of 100 articles from the previously defined database were examined, all published between 2020 and 2025 and distributed across high-impact journals. The corpus revealed a diversity of empirical approaches but converged around central concerns with accountability, transparency, risk, technological sovereignty, and global AI governance. The results are presented according to the three pillars of the AAP: Innovation and Regulation (Pillar I), Infrastructure and Security (Pillar II), and Diplomacy and Geopolitics (Pillar III). The scientific literature is broad and largely convergent around central concerns. Given the scope of the corpus analyzed, this section highlights only the most representative works in each topic, which serve as references for the critical discussion. The remaining studies were considered in the occurrence counts presented in the frequency tables.

4.1 Pillar I – Innovation and Regulation

Pillar I of the AAP proposes accelerating AI innovation through deregulation, promotion of open models, protection of freedom of expression, expansion of sectoral adoption, and workforce reskilling. In the analysis of the articles, this pillar was the most frequently addressed, with 345 occurrences distributed across the five categories.

Several studies addressed the tension between innovation and regulation. Chan et al. (2024) note that command-and-control approaches may hinder innovation, yet the absence of proportionate regulation increases risks of social exclusion and corporate capture. Kaminski (2023) emphasizes that clear and adaptive regulations are fundamental to ensuring accountability, countering the deregulatory bias of the AAP. Of the articles analyzed, 72 explicitly discussed deregulation, with most critical of broadly liberal models. Thus, the AAP diverges from the academic consensus, which values risk-based regulatory calibration.

The promotion of open-source and open-weight models was widely addressed in the literature, with 68 occurrences. Bommasani et al. (2024) highlight that open models enhance scientific innovation but create security and misuse risks. Koshiyama et al. (2024) reinforce the need for technical audits and standardization frameworks, partially converging with the AAP, which emphasizes openness but without sufficient attention to accountability.

The literature reiterates the centrality of fundamental rights such as non-discrimination, transparency, and diversity (Bélisle-Pipon et al., 2023; Wirtz et al., 2020). Studies such as Yoganathan et al. (2025) show that citizens tend to prefer government regulation when they perceive accountability risks. By prioritizing freedom of expression and minimizing DEI, the AAP departs from this consensus, raising concerns about social and international legitimacy.

The adoption of AI in governments and critical sectors was analyzed in 65 articles. Chen et al. (2024) demonstrate that the use of chatbots in U.S. state governments requires continuous evaluation and regulatory sandboxes. Bell et al. (2023) suggest guidelines for algorithmic transparency in public procurement. The literature converges with the AAP on the need for government adoption but warns of shortcomings in the absence of impact assessments and oversight mechanisms.

The impact of AI on labor appeared in 56 articles. Albous et al. (2025) show that national strategies must include robust training policies to address technological unemployment risks. Papyshev and Yarime (2023) reinforce that lack of attention to

inequalities may exacerbate social asymmetries. The AAP mentions reskilling but does not detail strategies for mitigating inequalities, which distances it from the literature. Overall, the literature recognizes the importance of accelerating innovation but advocates risk-based regulation, audits, and social safeguards. The AAP only partially converges, with a strong emphasis on deregulation and less attention to rights and equity.

4.2 Pillar II – Infrastructure and Security

Pillar II of the AAP focuses on building energy and technological infrastructure, restoring semiconductor manufacturing, and strengthening cybersecurity. This pillar ranked second in frequency, with 198 occurrences in the articles. A total of 71 occurrences related to digital security and resilience were identified. Ferretti (2022) argues that only the state can ensure AI security in critical sectors. Koshiyama et al. (2024) highlight the need for continuous audits and secure-by-design approaches. The AAP is largely aligned with this point, providing secure military data centers and incident response capacity, consistent with the literature.

The promotion of applied research and the use of AI to accelerate scientific discovery appeared in 45 articles. Radu (2021) and Robinson (2025) emphasize the importance of research infrastructure for responsible innovation. The AAP mentions support for research but focuses instrumentally on defense and security, less aligned with the multidimensional perspective described in the literature. In summary, there is strong convergence between the AAP and the literature regarding technological sovereignty and cybersecurity, but divergence regarding the integration of environmental criteria and the conception of research as a public good.

4.3 Pillar III – Diplomacy and Geopolitics

Pillar III of the AAP seeks to globally disseminate the American technological stack and counterbalance China's influence, in addition to reinforcing export controls and biosecurity measures. Although the least frequent in the corpus, it remained relevant, with 142 occurrences. The export of technologies and standards was addressed in 53 articles. Ingersleben-Seip (2023) highlights the strategic role of defining international standards. Albous et al. (2025) show how countries seek adherence to global standards to attract

investment. The AAP advocates unilateral diffusion, while the literature emphasizes multilateral cooperation.

The U.S. and China AI dispute was analyzed in 39 articles. Radu (2021) describes global governance as a space of both competition and cooperation, while (Robinson, 2025) suggests the creation of an international AI agency. The AAP acknowledges the competition but does not contemplate cooperative mechanisms, diverging from literature. A total of 28 occurrences related to export controls were found. Roberts et al. (2023) argue for regulatory harmonization rather than unilateralism. The AAP converges on the need for control but adopts a sovereigntist perspective. The literature mentioned 22 cases of emerging risk mitigation, especially in biotechnology (Bélisle-Pipon et al., 2023; Wörsdörfer, 2024). The AAP recognizes this challenge but focuses on national security, while the literature stresses transparency and social participation. The literature thus values international cooperation, harmonization, and multilateral governance, while the AAP privileges unilateralism and the projection of American supremacy. See Table 4.

Table 4

Occurrence Counts

Pillar	Category	Occurrences (n)	% of total
Pillar I – Innovation and Regulation			
	C1 – Deregulation	72	10.50%
	C2 – Open models and standards	68	9.90%
	C3 – Protection of values and rights	84	12.30%
	C4 – Sectoral and public adoption	65	9.50%
	C5 – Labor and workforce training	56	8.20%
	Subtotal Pillar I	**345**	**50.40%**
Pillar II – Infrastructure and Security			
	C6 – Technological infrastructure	82	12.00%
	C7 – Cybersecurity and resilience	71	10.40%
	C8 – Research and applied science	45	6.60%
	Subtotal Pillar II	**198**	**28.90%**
Pillar III – Diplomacy and Geopolitics			
	C9 – Export and global standards	53	7.70%
	C10 – Strategic competition	39	5.70%
	C11 – Export controls	28	4.10%
	C12 – Biosecurity and emerging risks	22	3.20%
	Subtotal Pillar III	**142**	**20.70%**
	TOTAL	**685**	**100%**

Source: Elaborated by the authors.

The content analysis demonstrated that the 100 articles are relatively evenly distributed across the three pillars, with greater emphasis on innovation/regulation (Pillar I). The results reveal convergences (cybersecurity, technological sovereignty, standards) and divergences (excessive emphasis on deregulation, rejection of DEI and climate concerns, diplomatic unilateralism).

5. DISCUSSION

The analysis indicates that the AAP occupies a distinctive position among national AI policies, both for its breadth and for the regulatory inflection it advances. By prioritizing deregulation, infrastructure expansion, and diplomatic unilateralism, the AAP signals a strategic repositioning of the United States amid global competition.

Pillar I of the AAP maintains that AI innovation should be promoted by removing regulatory barriers and revisiting "onerous" frameworks. This emphasis on deregulation stands in sharp contrast to the literature, which underscores the need for risk-based, proportionate regulation and audit mechanisms (Kaminski, 2023; Koshiyama et al., 2024; Wörsdörfer, 2024). The analysis shows that more than half of the examined articles (50.4%) highlight accountability, transparency, and fundamental rights as indispensable conditions for socially legitimate innovation.

Whereas the AAP elevates freedom of expression as the guiding principle, literature stresses the integration of values such as diversity, equity, and inclusion (Bélisle-Pipon et al., 2023; Wirtz et al., 2020). This divergence suggests that a unilateral emphasis on a single value may compromise social legitimacy and impede international acceptance of the U.S. model. Moreover, empirical studies show that citizens tend to prefer state regulation when they perceive risks tied to a lack of corporate responsibility (Wilczek et al., 2025; Yoganathan et al., 2025), reinforcing the critique that the American strategy may face domestic and external social resistance.

In Pillar II, which addresses technological infrastructure, semiconductors, and cybersecurity, convergence between the AAP and the literature is stronger. Both recognize the importance of ensuring technological sovereignty, securing resilient supply chains, and strengthening the cybersecurity of critical systems (Ferretti, 2022; Russell &

McGravey, 2025). Alignment is particularly clear in support for secure-by-design approaches and in the valuation of semiconductors as a strategic resource.

Nonetheless, important gaps remain. Recent literature warns of environmental and sustainability impacts associated with large-scale AI infrastructure, including the energy intensity of data centers and model training (Han et al., 2025; Koenig et al., 2023). The AAP explicitly rejects integrating the climate agenda into AI policy, a mismatch with emerging international practice. This omission may diminish the plan's global legitimacy, especially in multilateral fora where sustainability is a central evaluative criterion.

Another point of divergence concerns the financing and use of scientific research. While literature emphasizes research as a public good and an instrument of international cooperation (Radu, 2021; Robinson, 2025), the AAP prioritizes applied research for military and defense objectives. This choice narrows the scope of science and may limit international collaboration, reinforcing the perception that the U.S. strategy instrumentalizes AI primarily for national security.

Pillar III represents the AAP's most critical dimension when viewed against literature. The plan bets on the unilateral export of the "American technological stack," on counterbalancing China's influence, and on strengthening semiconductor export controls. Although literature acknowledges geopolitical competition as a structural element of global AI governance (Ingersleben-Seip, 2023; Radu, 2021), it emphasizes that the sustainability of the international regime depends on multilateral cooperation and jointly defined standards (Albous et al., 2025; Roberts et al., 2023).

The data show that only 20.7% of corpus occurrences relate to Pillar III, suggesting that the literature still devotes less attention to technological diplomacy than to domestic regulation and infrastructure. Even so, existing studies indicate a clear trend: AI governance should be built cooperatively, with regulatory harmonization and the establishment of international agencies (Robinson, 2025). By privileging unilateralism and sovereignty, the AAP risks amplifying global regulatory fragmentation and weakening the convergence needed to address cross-border risks such as disinformation, biotechnology, and cybersecurity. Overall, the discussion indicates that the AAP produces selective convergences and structural divergences vis-à-vis the scientific literature. It converges on cybersecurity, technological sovereignty, and the need for technical standards, yet diverges on core aspects such as accountability, equity, sustainability, and international cooperation.

This landscape reveals a trade-off: while the AAP may reinforce U.S. competitiveness in the short term through deregulation and technological nationalism, it may also erode the United States' international legitimacy in AI governance leadership. By not integrating values widely recognized in the literature – such as inclusion, transparency, sustainability, and multilateralism – the plan risks consolidating a hegemonic yet weakly cooperative model that is susceptible to social and diplomatic resistance.

Another trade-off concerns the balance between innovation speed and accountability capacity. Pillar I of the AAP prioritizes rapid diffusion of AI through deregulation, promotion of open models, and streamlined compliance, framing rules perceived as "onerous" as barriers to competitiveness. This orientation collides with the dominant pattern in the corpus, in which risk-based calibration, audits, impact assessment, and enforceable accountability instruments are treated as preconditions for legitimate innovation. Empirically, this tension is visible in the frequency profile of Pillar I, where deregulation appears in 72 occurrences and open models in 68, alongside an even higher concentration of discussions on protection of values and rights (84 occurrences), indicating that innovation agendas are repeatedly debated in the same analytical space as safeguards, transparency, and non-discrimination. In practical terms, the corpus associates innovation speed with governance instruments that preserve verifiability, such as regulatory sandboxes coupled with continuous evaluation, structured auditing regimes, and procurement transparency obligations. The AAP partially converges on experimentation and scaling, but the evidence synthesized in the review suggests that absent robust auditability and institutional oversight, acceleration strategies can shift compliance from a substantive to a procedural register, increasing exposure to legitimacy erosion and implementation gaps even when short-term competitiveness improves.

The third trade-off concerns infrastructure expansion versus environmental sustainability and long-run feasibility. Pillar II frames AI supremacy as dependent on large-scale investments in energy, semiconductors, and data centers, including simplification of environmental licensing and expansion of computer-intensive infrastructures. Literature, however, treats sustainability not as a peripheral add-on, but as a policy constraint that reshapes infrastructure choices, affects international acceptance, and conditions the durability of governance regimes through the management of environmental externalities. This divergence is explicit in the AAP assessment, which

registers that the plan boosts chips and data centers while rejecting climate-related agendas, generating partial alignment at best with scholarship that expects sustainability criteria to be embedded in infrastructure-centric sovereignty strategies. The tension is therefore not merely normative: it has operational consequences in licensing design, grid and energy planning, data center sitting, and the governance of computation-related externalities. In the framework advanced by the article, sustainable infrastructure policy requires feedback loops that recognize ecological impacts as governance-relevant variables; otherwise, infrastructure resilience and sovereignty objectives risk being achieved through short-cycle scale-up that undermines long-run legitimacy and international convergence pressures, particularly when peer jurisdiction incorporate sustainability into standards, procurement, and accountability requirements.

From a theoretical standpoint, the findings reinforce the importance of understanding AI governance not only as an instrument of industrial policy but as a global normative regime grounded in trust, cooperation, and regulatory convergence (Bannister & Connolly, 2020; Papyshev & Yarime, 2023). From a practical standpoint, the study suggests that countries seeking to align with the U.S. model must weigh the costs of legitimizing a unilateral arrangement over more inclusive multilateral practice. Table 5 consolidates the AAP's assessment vis-à-vis the scientific literature.

Table 5

AAP Assessment

Pillar / Category	Scientific literature and reference	AAP position	Classif.
Pillar I - Innovation and Regulation			
C1: Deregulation	Risk-based, proportionate regulation and audits	Emphasis on removing regulatory barriers and revising "onerous" rules	Misaligned
C2: Open models and standards	Encourage open innovation with safeguards and technical audits	Promotion of open-source and open-weight with limited emphasis on accountability	Partially aligned
C3: Protection of values and rights	Centrality of DEI, transparency, and non-discrimination	Revision of NIST AI RMF removing DEI and climate references; prioritizes freedom of expression	Misaligned
C4: Sectoral and public adoption	Government use with sandboxes, impact assessments, and ethical procurement	Encourages AI in health, government, and defense with limited safeguard detail	Partially aligned
C5: Labor and skills	Reskilling strategies and mitigation of inequalities	Provides training and reskilling, with limited	Partially aligned

		focus on structural inequalities	
Pillar II - Infrastructure and Security			
C6: Technological infrastructure	Digital sovereignty with sustainability criteria	Boosts chips and data centers while explicitly rejecting the climate agenda	Partially aligned
C7: Cybersecurity and resilience	Secure-by-design, continuous audits, and critical-system protection	Secure military data centers and incident response capacity	Aligned
C8: Research and applied science	Research as a public good and vector for international cooperation	Support for applied research aimed at defense and national competitiveness	Misaligned
Pillar III - Diplomacy and Geopolitics			
C9: Export and global standards	Multilateral cooperation and standards convergence	Unilateral export of the "American stack"	Misaligned
C10: Strategic competition	Global governance as a space for competition and cooperation	Emphasis on competition with China and unilateral leadership	Partially aligned
C11: Export controls	Regulatory harmonization to avoid fragmentation	Reinforced export controls and chip localization verification	Misaligned
C12: Biosecurity and emerging risks	Mitigation via audits, transparency, and cooperation	Recognizes frontier risks, framing primarily within national security	Partially aligned

Source: Elaborated by the authors.

Table 5 synthesizes the critical assessment of the AAP against the scientific literature, highlighting selective convergences and structural divergences. The plan is aligned only in specific areas such as cybersecurity (C7); it is partially aligned on open models, labor, infrastructure, and biosecurity (C2, C4, C5, C6, C10, C12); and misaligned on core areas including deregulation, rights protection, scientific research, and multilateral diplomacy (C1, C3, C8, C9, C11). Overall, while the AAP advances innovation and technological sovereignty, it diverges from international best practices by prioritizing broad deregulation and unilateralism and by foregrounding national security at the expense of inclusion, cooperation, and sustainability.

A broader theoretical synthesis emerges from these findings, transcending the empirical boundaries of the AAP and delineating a conceptual model for understanding how artificial intelligence governance evolves as a multidimensional regulatory regime. Rather than being limited to the assessment of a single national strategy, this framework integrates the main constructs identified in the literature into a cohesive theoretical

structure. In this sense, AI governance can be interpreted as a complex adaptive system shaped by interdependent dimensions that balance technological progress, institutional legitimacy, and normative alignment. The literature reviewed indicates that legitimacy in AI governance does not arise solely from technical efficiency or regulatory control, but from the dynamic equilibrium between innovative incentives, regulatory safeguards, and transnational coordination mechanisms.

Building on this perspective, the following theoretical framework abstracts the core mechanisms underpinning contemporary AI governance (Table 6). It conceptualizes this field as a multi-level and multi-actor regime, where governments, private organizations, and epistemic communities interact to co-produce standards, distribute authority, and negotiate values. This model provides an analytical basis for comparing national and international strategies, identifying the key variables and causal linkages that explain how legitimacy, cooperation, and competitiveness coexist within emerging regimes of artificial intelligence governance.

Table 6

Theoretical framework: multi-dimensional regime of AI governance

Analytical Dimension	Core Theoretical Constructs	Underlying Logics and Mechanisms	Expected Governance Outcomes	Operational examples	References
1. Innovation and Regulatory Calibration	Balancing innovation incentives with regulatory control; adaptive governance; proportional risk-based regulation	Dynamic equilibrium between technological progress and normative constraint through iterative feedback between policy, market, and ethics	Regulatory proportionality, accountability, and sustained innovation legitimacy	Risk-tiering rules in licensing and oversight; regulatory sandboxes with performance and safety thresholds; mandatory algorithmic impact assessments for high-risk uses	Bannister & Connolly (2020); Kaminski (2023); Wirtz et al. (2020)
2. Technological Sovereignty and Infrastructure Resilience	National capability in digital infrastructure; strategic autonomy; secure-by-design systems	Interdependence between physical (hardware) and institutional (policy) infrastructures; resilience as a condition for technological independence	Strengthened digital sovereignty and resilience of critical infrastructures	Public procurement requirements for secure-by-design systems; national compute strategies and semiconductor supply safeguards; resilience standards for critical digital infrastructure and cybersecurity baselines	Ferretti (2022); Russell & McGravey (2025)
3. Ethical and Social Legitimacy	Inclusion, transparency, and fairness as pillars of algorithmic trust	Legitimation through procedural fairness and value pluralism in design, deployment, and evaluation of AI systems	Social acceptance, reduction of algorithmic bias, and trust in AI governance	Transparency notices and explainability obligations in public services; bias audits and fairness testing protocols; participatory governance mechanisms for affected communities	Bélisle-Pipon et al. (2023); Wirtz et al. (2020); Yoganathan et al. (2025)
4. Sustainability and Environmental Responsibility	Integration of environmental externalities into AI governance frameworks	Feedback between digital transformation and ecological impacts; “green AI” paradigm	Low-carbon infrastructure, responsible computation, and long-term ecological alignment	Environmental criteria for data center licensing and public procurement; energy efficiency targets for compute-intensive workloads; reporting of carbon footprint for large-scale training and inference	Han et al. (2025); Koenig et al. (2023)
5. Multilevel and Cooperative Governance	Coordination among national, transnational, and	Governance as a distributed network of norms and standards negotiated across jurisdictions	Convergence of standards, interoperability, and	Alignment with internationally recognized risk management frameworks; mutual recognition mechanisms for conformity assessments;	Radu (2021); Roberts et al. (2023); Robinson (2025)

	private actors; polycentric regulation		legitimacy in global regimes	cross-border cooperation arrangements for incident reporting and enforcement	
6. Geopolitical Dynamics and Regime Competition	Interaction between hegemonic models of AI governance (U.S., EU, China)	Regulatory competition as a driver of diffusion, mimicry, and divergence in global norms	Hybrid regimes balancing sovereignty with international alignment	Export controls and strategic investment screening for AI-critical assets; regulatory benchmarking and "race to standards" initiatives; diffusion of governance models through trade and security partnerships	Albous et al. (2025); Ingersleben-Seip (2023)
7. Knowledge Co-Production and Epistemic Communities	Integration of scientific, political, and societal expertise in AI policy design	Epistemic pluralism as a mechanism for accountability and learning	Collective intelligence and evidence-based policy innovation	Independent scientific advisory boards for AI policy; structured stakeholder consultations in rulemaking; public registries of AI systems and evidence repositories supporting impact evaluation	Papyshev & Yarime (2023); Djeffal et al. (2022)

Source: Elaborated by the authors.

The theoretical framework presented (Table 6) systematizes the central analytical dimensions that underpin the contemporary governance of artificial intelligence. It synthesizes insights from the literature into a multi-dimensional model, emphasizing that AI governance operates through the interaction of technological, normative, institutional, and geopolitical forces. Each dimension, such as innovation, sovereignty, ethics, sustainability, cooperation, geopolitics, and knowledge co-production, represents a specific mechanism through which legitimacy and stability are constructed within governance regimes.

In this configuration, governance effectiveness depends on the system's ability to maintain equilibrium among competing logics: promoting innovation without eroding accountability, strengthening sovereignty without undermining cooperation, and advancing security without sacrificing inclusion or sustainability. At the same time, these trade-offs unfold under persistent informational asymmetries between public authorities, dominant technology firms, and epistemic communities, which condition problem definition, instrument choice, and the verifiability of compliance. Under such asymmetries, the risk of regulatory capture increases, particularly in audit regimes, standard-setting processes, and self-regulatory arrangements, with direct implications for both legitimacy and enforcement capacity. Legitimate and sustainable AI governance therefore requires adaptive coordination between regulatory flexibility, infrastructural resilience, and normative convergence, while institutionalizing safeguards that reduce informational dependence and preserve the autonomy of oversight.

In contemporary technology governance, informational asymmetry is not a peripheral issue but a structural condition that shapes how states define problems, select instruments, and enforce compliance. When regulatory questions are complex and technically mediated, decision makers often depend on expert knowledge networks to interpret causal mechanisms, define what counts as "risk," and translate uncertainty into actionable policy choices (Haas, 1992). This dependence creates a persistent imbalance between public authorities and regulated actors that control key resources, including data, computing, proprietary technical documentation, and the practical know-how required to operationalize audits and safety assessments. Classic political economy accounts show how such asymmetries expand the discretion of intermediaries and amplify the leverage of information holders, affecting both the design of regulatory incentives and the monitoring of outcomes (Dal Bó, 2006; Laffont & Tirole, 1991). Therefore, AI

governance regimes tend to evolve under conditions in which the state's capacity to specify, verify, and sanction is partially contingent on cooperation from the very actors it seeks to regulate, which heightens the relevance of accountability arrangements and the institutional architecture of oversight.

These conditions also increase exposure to regulatory capture, especially in governance modalities that rely on auditability, standard setting, and voluntary commitments. Capture risks emerge not only through overt lobbying but also through subtler mechanisms such as "informational capture," revolving doors, and the gradual alignment of regulatory priorities with the problem framings and technical vocabularies promoted by dominant firms and industry aligned expert communities (Carpenter & Moss, 2013; Dal Bó, 2006). In AI, where transparency is frequently treated as the main route to accountability, scholarship has shown that "seeing" a system does not necessarily translate into knowing it sufficiently to govern it, particularly when access is partial, selectively disclosed, or constrained by proprietary protections (Ananny & Crawford, 2018). Likewise, the institutionalization of algorithmic auditing can be undermined when internal audit processes and documentation remain firm controlled or when compliance becomes a procedural exercise that prioritizes innovation speed and market positioning over substantive rights protection (Raji et al., 2020). Recent contributions explicitly connect AI governance and capture to strategic competition and regulatory arbitrage, arguing that high stakes geopolitical rivalry can reinforce incentives to trade off stringent accountability and sustainability requirements in favor of innovation and industrial leadership, with downstream effects on legitimacy and international cooperation (Lancieri et al., 2025; Metcalf, 2025; Veale et al., 2023). This framing directly informs the tradeoffs explored in the analysis, particularly tensions between innovation and accountability, scale up of infrastructure and environmental responsibility, and sovereignty-oriented approaches versus multilateral coordination.

By abstracting these relationships, the model advances a general theory of AI governance that transcends policy cases. It provides conceptual grounding for future comparative analyses, enabling the identification of patterns of convergence and divergence across national and international contexts. Ultimately, the framework contributes to positioning AI governance as a global normative regime, where legitimacy emerges not only from state authority or technological capability but from the collective alignment of innovation, ethics, and cooperation.

6. CONCLUSION

The research sought to answer the question: How can the AAP be evaluated considering best practices in artificial intelligence regulation and governance identified in scientific literature? The main objective was to critically analyze the AAP, identifying consistencies, gaps, and trade-offs across its three pillars: innovation, infrastructure, and diplomacy, as well as to assess its alignment with global regulatory trends.

The analysis, based on an integrative review of 100 high-impact articles, revealed selective convergences and structural divergences between the AAP and international academic consensus. Convergences occur mainly in the areas of cybersecurity, technological sovereignty, and standardization, while divergences emerge in relation to accountability, diversity, sustainability, and multilateral cooperation. The study demonstrates that although the AAP strengthens the United States' technological competitiveness, it also reinforces unilateralism and deregulation, which may weaken global legitimacy and the cooperative basis required for sustainable AI governance.

From a theoretical perspective, the research contributes to consolidating the understanding of AI governance as a multidimensional and adaptive regime, in which legitimacy derives from the balance between innovation, regulation, and international coordination. The proposed framework advances this view by integrating analytical dimensions such as sovereignty, ethics, sustainability, cooperation, and epistemic co-production, offering a broader conceptual lens for interpreting the dynamics of global AI governance. At the practical level, the findings suggest that policymakers should adopt adaptive regulation capable of reconciling technological dynamism with ethical responsibility. This includes promoting risk-based mechanisms, institutional transparency, and multilevel coordination to reduce fragmentation in global governance systems. Socially, the study highlights the importance of embedding inclusion, diversity, and sustainability within AI policies, ensuring that technological progress aligns with collective well-being and democratic values. The absence of these dimensions may deepen inequalities and erode public trust in AI systems.

This study has limitations that should be made explicit when interpreting the results. The exclusive reliance on Scopus and English-language, peer-reviewed publications may introduce coverage, language, and publication biases, potentially

underrepresenting region-specific debates, policy-oriented outlets, and relevant scholarship indexed unevenly across citation databases. Evidence also indicates that citation indices differ materially in backward and forward citation coverage and accuracy, which can affect what is discoverable and, indirectly, what appears “central” in a synthesis (Gusenbauer, 2024). While the snowballing extension helps mitigate retrieval blind spots by leveraging citation networks beyond the initial query set (Wohlin, 2014), it does not fully eliminate structural biases associated with database scope and English-only filters.

Policymakers and regulators should treat governance effectiveness as a balance between innovation enablement and enforceable accountability, institutionalizing mandatory instruments such as impact assessments, auditability requirements, and transparency obligations for high-risk uses, while using regulatory sandboxes as controlled learning infrastructures with clear entry and exit criteria and continuous evaluation. In parallel, infrastructure and sovereignty strategies should embed sustainability as a governance constraint by integrating environmental performance criteria into procurement and licensing decisions for data centers and compute-intensive projects, since literature associates long-run feasibility and legitimacy with explicit management of environmental externalities. Finally, reducing fragmentation requires convergence-ready standard setting and institutional participation in multilateral venues, combined with safeguards that reduce informational dependence on dominant firms and mitigate capture risks in audit and standard-setting processes through independent expertise, conflict-of-interest controls, and verifiable documentation.

Future research should extend this study through comparative designs that examine how the AAP aligns or diverges from other governance models, especially the EU and China, and include additional regions to test whether observed patterns are regime-specific or structurally recurrent. Longitudinal approaches are also needed to track how regulatory priorities, standards capacity, audit regimes, and enforcement infrastructures evolve over time under policy feedback and geopolitical pressures, while interviews and process tracing with policymakers, regulators, and standards bodies can clarify how trade-offs are negotiated in practice and how informational asymmetries shape instrument choice and implementation feasibility. Mixed-method strategies combining document analysis, network mapping of standard diffusion, and empirical measurement of legitimacy outcomes, such as trust and perceived fairness, would strengthen causal explanations and improve the evaluation of governance effectiveness

across contexts. There are needs to expand the comparative scope to include the EU AI Act, the Brazilian Artificial Intelligence Strategy, and Asian governance frameworks, incorporating quantitative bibliometric analyses and policy network mapping.

REFERENCES

Albous, M. R., Al-Jayyousi, O. R., & Stephens, M. (2025). AI Governance in the GCC States: A Comparative Analysis of National AI Strategies. *Journal of Artificial Intelligence Research*, *82*, 2389-2422. https://doi.org/10.1613/jair.1.17619

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, *20*(3), 973-989. https://doi.org/10.1177/1461444816676645

Bannister, F., & Connolly, R. (2020). Administration by algorithm: A risk management framework. *Information Polity*, *25*(4), 471-490. https://doi.org/10.3233/IP-200249

Bardin, L. (2008). *Análise De Conteúdo*. Edições 70.

Bélisle-Pipon, J.-C., Monteferrante, E., Roy, M.-C., & Couture, V. (2023). Artificial intelligence ethics has a black box problem. *AI & SOCIETY*, *38*(4), 1507-1522. https://doi.org/10.1007/s00146-021-01380-0

Bell, A., Nov, O., & Tandon, J. S. (2023). Think about the stakeholders first! Toward an algorithmic transparency playbook for regulatory compliance. *Data & Policy*, *5*, e12. https://doi.org/10.1017/dap.2023.8

Bommasani, R., Kapoor, S., Klyman, K., Longpre, S., Ramaswami, A., Zhang, D., Schaake, M., Ho, D. E., Narayanan, A., & Liang, P. (2024). Considerations for governing open foundation models. *Science*, *386*(6718), 151-153. https://doi.org/10.1126/science.adp1848

Carpenter, D., & Moss, D. A. (2013). Introduction. In D. Carpenter & D. A. Moss (Eds.), *Preventing regulatory capture* (pp. 1–22). Cambridge University Press. https://doi.org/10.1017/CBO9781139565875.002

Chan, K. J. D., Papyshev, G., & Yarime, M. (2024). Balancing the tradeoff between regulation and innovation for artificial intelligence: An analysis of top-down command and control and bottom-up self-regulatory approaches. *Technology in Society*, *79*, 102747. https://doi.org/10.1016/j.techsoc.2024.102747

Chen, T., Gascó-Hernandez, M., & Esteve, M. (2024). The Adoption and Implementation of Artificial Intelligence Chatbots in Public Organizations: Evidence from U.S. State Governments. *The American Review of Public Administration*, *54*(3), 255-270. https://doi.org/10.1177/02750740231200522

Dafoe, A. (2023). AI governance: Overview and theoretical lenses. In J. B. Bullock, Y.-C. Chen, J. Himmelreich, V. M. Hudson, A. Korinek, M. M. Young, & B. Zhang (Eds.), *The Oxford handbook of AI governance* (pp. 21-44). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.2

Dal Bó, E. (2006). Regulatory Capture: A Review. *Oxford Review of Economic Policy*, *22*(2), 203-225. https://doi.org/10.1093/oxrep/grj013

Djeffal, C., Siewert, M. B., & Wurster, S. (2022). Role of the state and responsibility in governing artificial intelligence: A comparative analysis of AI strategies. *Journal of European Public Policy*, *29*(11), 1799-1821. https://doi.org/10.1080/13501763.2022.2094987

Ferretti, T. (2022). An Institutionalist Approach to AI Ethics: Justifying the Priority of Government Regulation over Self-Regulation. *Moral Philosophy and Politics*, *9*(2), 239-265. https://doi.org/10.1515/mopp-2020-0056

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, *45*, 105681. https://doi.org/10.1016/j.clsr.2022.105681

Gusenbauer, M. (2024). Beyond Google Scholar, Scopus, and Web of Science: An evaluation of the backward and forward citation coverage of 59 databases' citation indices. *Research Synthesis Methods*, *15*(5), 802-817. https://doi.org/10.1002/jrsm.1729

Haas, P. M. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization*, *46*(1), 1-35. https://doi.org/10.1017/S0020818300001442

Han, S., Zhang, H., Li, H., & Xun, Z. (2025). Digital transformation and carbon emission reduction: The moderating effect of external pressure and support. *Journal of Cleaner Production*, *500*, 145108. https://doi.org/10.1016/j.jclepro.2025.145108

Ingersleben-Seip, N. von. (2023). Competition and cooperation in artificial intelligence standard setting: Explaining emergent patterns. *Review of Policy Research*, *40*(5), 781-810. https://doi.org/10.1111/ropr.12538

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Kaminski, M. E. (2023). Regulating the risks of AI. *Boston University Law Review*, *103*(5), 1347-1411.

Koenig, P. D., Wurster, S., & Siewert, M. B. (2023). Sustainability challenges of artificial intelligence and citizens' regulatory preferences. *Government Information Quarterly*, *40*(4), 101863. https://doi.org/10.1016/j.giq.2023.101863

Koshiyama, A., Kazim, E., Treleaven, P., Rai, P., Szpruch, L., Pavey, G., Ahamat, G., Leutner, F., Goebel, R., Knight, A., Adams, J., Hitrova, C., Barnett, J., Nachev, P., Barber, D., Chamorro-Premuzic, T., Klemmer, K., Gregorovic, M., Khan, S., … Chatterjee, S. (2024). Towards algorithm auditing: Managing legal, ethical and technological risks of AI, ML and associated algorithms. *Royal Society Open Science*, *11*(5), 230859. https://doi.org/10.1098/rsos.230859

Laffont, J.-J., & Tirole, J. (1991). The Politics of Government Decision-Making: A Theory of Regulatory Capture. *The Quarterly Journal of Economics*, *106*(4), 1089-1127. https://doi.org/10.2307/2937958

Lancieri, F., Edelson, L., & Bechtold, S. (2025). AI regulation: Competition, arbitrage and regulatory capture. *Theoretical Inquiries in Law*, *26*(1), 239-262. https://doi.org/10.1515/til-2025-0010

Lebeaux, R. (2025, August 14). *How America's 'AI Action Plan' May Help Shape AI Governance Globally*. https://deloitte.wsj.com/sustainable-business/how-americas-ai-action-plan-may-help-shape-ai-governance-globally-4b3a054e

Marchiori, S., Hopster, J. K. G., Puzio, A., Riemsdijk, M. B. V., Kraaijeveld, S. R., Lundgren, B., Viehoff, J., & Frank, L. E. (2025). A Social Disruptiveness-Based Approach to AI Governance: Complementing the Risk-Based Approach of the AI Act. *Science and Engineering Ethics*, *31*(5), 25. https://doi.org/10.1007/s11948-025-00545-0

Metcalf, T. (2025). AI safety and regulatory capture. *AI & SOCIETY*, 1-16. https://doi.org/10.1007/s00146-025-02534-0

Papyshev, G., & Yarime, M. (2023). The state's role in governing artificial intelligence: Development, control, and promotion through national strategies. *Policy Design and Practice*, *6*(1), 79-102. https://doi.org/10.1080/25741292.2022.2162252

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178-193. https://doi.org/10.1080/14494035.2021.1929728

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020, January). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, Barcelona, Spain. https://doi.org/10.1145/3351095.3372873

Roberts, H., Babuta, A., Morley, J., Thomas, C., Taddeo, M., & Floridi, L. (2023). Artificial intelligence regulation in the United Kingdom: A path to good governance and global leadership? *Internet Policy Review*, *12*(2), 1-31. https://doi.org/10.14763/2023.2.1709

Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: Barriers and pathways forward. *International Affairs*, *100*(3), 1275-1286. https://doi.org/10.1093/ia/iiae073

Robinson, M. (2025). The establishment of an international AI agency: An applied solution to global AI governance. *International Affairs*, *101*(4), 1483-1497. https://doi.org/10.1093/ia/iiaf105

Russell, A. L., & McGravey, K. (2025). Digital oil: Chips, artificial intelligence and US national security. *International Affairs*, *101*(3), 1087-1101. https://doi.org/10.1093/ia/iiaf009

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1)*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

The White House. (2025). *Winning the Race America's AI Action Plan*. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

Veale, M., Matus, K., & Gorwa, R. (2023). AI and Global Governance: Modalities, Rationales, Tensions. *Annual Review of Law and Social Science*, *19*(1), 255-275. https://doi.org/10.1146/annurev-lawsocsci-020223-040749

Wilczek, B., Thaesler-Kordonouri, S., & Eder, M. (2025). Government regulation or industry self-regulation of AI? Investigating the relationships between uncertainty avoidance, people's AI risk perceptions, and their regulatory preferences in Europe. *AI & SOCIETY*, *40*(5), 3797-3811. https://doi.org/10.1007/s00146-024-02138-0

Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, *43*(9), 818-829. https://doi.org/10.1080/01900692.2020.1749851

Wohlin, C. (2014, June). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference*

on Evaluation and Assessment in Software Engineering, London, England. https://doi.org/10.1145/2601248.2601268

Wörsdörfer, M. (2024). Mitigating the adverse effects of AI with the European Union's artificial intelligence act: Hype or hope? *Global Business and Organizational Excellence*, *43*(3), 106-126. https://doi.org/10.1002/joe.22238

Yoganathan, V., Osburg, V.-S., & Janakiraman, N. (2025). Lending Legitimacy to Corporate Digital Responsibility: Trust in Firm Versus Government Regulation of Artificial Intelligence Services. *Journal of Service Research*, 1-21. https://doi.org/10.1177/10946705251345097

Zotero. (2025). *IA regulation*. https://www.zotero.org/groups/6279590/ia_regulation

Darci de Borba Santos Júnior

Ph.D. in Business Administration from the University of Vale do Rio dos Sinos (Unisinos); Master's in Strategic Administration from the Pontifical Catholic University of Rio Grande do Sul (PUCRS); Researcher at the Institute for Applied Economic Research (Ipea). E-mail: darci.junior@ipea.gov.br

Rafael Brinkhues

Doctor in Business Administration from the Federal University of Rio Grande do Sul (UFRGS), Master in Business Administration from the Pontifical Catholic University of Paraná (PUCPR) and from the Federal University of Minas Gerais (UFMG); Full Professor at the Federal Institute of Education, Science and Technology of Rio Grande do Sul (IFRS) and Project Manager at the Presidency of the Federative Republic of Brazil. E-mail: rafael.brikhues@presidencia.gov.br

AUTHORS' CONTRIBUTION STATEMENT

Darci de Borba Santos Júnior: Conceptualization (Lead); Data curation (Lead); Formal analysis (Lead); Investigation (Lead); Methodology (Lead); Project administration (Lead); Resources (Lead); Supervision (Lead); Validation (Equal); Visualization (Equal); Writing, original draft (Lead); Writing, review & editing (Equal).

Rafael Alfonso Brinkhues: Conceptualization (Supporting); Data curation (Supporting); Formal analysis (Supporting); Investigation (Supporting); Methodology (Supporting); Resources (Supporting); Supervision (Equal); Validation (Equal); Visualization (Equal); Writing, original draft (Supporting); Writing, review & editing (Equal).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

RESEARCH DATA AVAILABILITY STATEMENT

The entire dataset supporting the results of this study is publicly available in the Zotero repository. The collection includes all articles retrieved, screened, and analyzed during the integrative literature review. The dataset can be accessed at https://www.zotero.org/groups/6279590/ia_regulation

AI USAGE STATEMENT

This article was originally written in Portuguese. Artificial intelligence-based tools were used exclusively to support the translation and language revision process. All ideas, arguments, and analyses presented herein are entirely those of the author.

This preprint was submitted under the following conditions:

- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a Creative Commons CC-BY license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.