THE ENFORCEMENT GAP: WHY TRADITIONAL PRIVACY IMPACT ASSESSMENTS FAIL FOR AI SYSTEMS AND WHAT SHOULD REPLACE THEM

Amit Kumar Padhy, Research Scholar at National Law University, Nagpur, India

ABSTRACT

This article examines the critical enforcement gap between traditional Privacy Impact Assessments (PIAs) and the governance needs posed by contemporary artificial intelligence (AI) systems. While PIAs, particularly as codified under frameworks like the EU General Data Protection Regulation (GDPR), represent a foundational tool for privacy risk management, their static, linear, and data-centric design fundamentally fails to address the dynamic, opaque, and multifaceted risks presented by AI technologies. Through detailed legal and technical analysis, combined with empirical evidence of AI-related privacy incidents and enforcement challenges, this study demonstrates how traditional PIAs inadequately capture emergent algorithmic harms, including bias, opacity, and collective societal risks. The article critically evaluates emerging alternatives such as Algorithmic Impact Assessments (AIAs) and Fundamental Rights Impact Assessments (FRIAs), highlighting their enhanced scope but also their continued reliance on static assessment paradigms. Arguing that incremental reforms are insufficient, the article advocates for a reconceptualized AI governance framework grounded in adaptive, outcome-focused, and participatory accountability mechanisms that integrate continuous technical monitoring and robust institutional oversight. This framework aims to bridge the enforcement gap by aligning regulatory approaches with the evolving realities of AI systems, thereby safeguarding fundamental rights and restoring public trust in AI governance.

Keywords: AI Governance, Accountability, Right to Explanation, Data Privacy.

I. Introduction

The proliferation of artificial intelligence systems across virtually every sector of the economy has exposed a fundamental misalignment between existing data protection frameworks and the realities of algorithmic governance. Traditional Privacy Impact Assessments (PIAs)[1], enshrined in regulatory instruments such as Article 35 of the General Data Protection Regulation (GDPR), were designed for a world of static data processing operations with predictable inputs, linear workflows, and deterministic outputs.[2] Yet AI systems, particularly machine learning models and generative AI applications, operate through dynamic, probabilistic processes that evolve continuously through training, fine-tuning, and real-world deployment.[3]

This enforcement gap has become increasingly pronounced as AI incidents have surged by 56.4% in a single year, with 233 documented cases in 2024 alone. The regulatory response has been fragmented and reactive, with traditional data protection authorities struggling to apply decades-old assessment frameworks to technologies that fundamentally challenge core privacy principles such as purpose limitation, data minimization, and transparency.[4] The result is a governance vacuum where organizations deploy AI systems with significant privacy implications while nominally complying with existing PIA requirements that fail to capture the true scope of risks to individual rights and freedoms.[5]

The stakes of this misalignment extend far beyond technical compliance. Public trust in AI companies has declined from 50% to 47% between 2023 and 2024, while regulatory activity has more than doubled across 75 countries. The European Union's AI Act represents the most comprehensive attempt to address these challenges through novel instruments such as Fundamental Rights Impact Assessments (FRIAs), yet even these emerging frameworks remain anchored to traditional assessment methodologies that may prove inadequate for the unique characteristics of AI systems.

This Article argues that the enforcement gap between traditional PIAs and AI governance

[1] European Data Protection Board, "Guidelines on Data Protection Impact Assessment (DPIA)," EDPB, 2022.
[2] Regulation (EU) 2016/679, General Data Protection Regulation, Article 35.
[3] Kaminski, M.E. & Malgieri, G. "Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations," International Data Privacy Law, 11(2), 2021, pp. 125–144.
[4] Privacy Act of 1974, 5 U.S.C. § 552a (USA).
[5] Stanford AI Index Report 2025: "AI Data Privacy Wake-Up Call: Findings From Stanford's 2025 AI Index," Kiteworks, 2025.

demands not merely incremental reforms but a fundamental reconceptualization of algorithmic accountability frameworks. Through empirical analysis of PIA failures in AI contexts, comparative examination of emerging alternatives, and synthesis of technical and legal scholarship, this Article demonstrates that effective AI governance requires moving beyond static, compliance-oriented assessments toward dynamic, outcome-focused accountability mechanisms that can adapt to the evolving nature of AI systems while maintaining meaningful protection for fundamental rights.

II. Traditional PIAs: Framework and Limitations

The contemporary framework for Privacy Impact Assessments emerged from decades of evolution in data protection law, reaching its most sophisticated expression in the GDPR's Data Protection Impact Assessment (DPIA) requirements. Article 35 mandates DPIAs for processing operations "likely to result in a high risk to the rights and freedoms of natural persons," particularly those involving systematic and extensive evaluation of personal aspects through automated processing.[6] This framework reflects core assumptions about data processing that made sense in the pre-AI era: that processing purposes could be clearly defined ex ante, that data flows could be mapped with precision, and that risks could be identified and mitigated through technical and organizational measures.

The traditional PIA methodology follows a structured, linear progression: scope definition, data flow mapping, risk identification, impact assessment, and mitigation planning. Organizations must describe the processing operation's purpose, assess its necessity and proportionality, identify potential privacy risks, and implement safeguards to address those risks.[7] The process assumes that these elements remain relatively stable throughout the system's lifecycle, with periodic reviews sufficient to address any changes in risk profile.[8]

However, this framework embeds several critical limitations that become acute when applied to AI systems. First, the requirement for "clear and specific" purpose specification conflicts with the exploratory nature of many AI applications, where the full range of potential insights and applications may not be apparent at the outset. Machine learning systems are often

[6] GDPR, Article 35, Recitals 84, 89, 90.
[7] European Data Protection Board, "Guidelines on DPIA," EDPB, 2022.
[8] UK Information Commissioner's Office, "Conducting a Data Protection Impact Assessment (DPIA)," 2021.

designed to discover patterns and correlations that humans cannot anticipate, making it impossible to fully specify their purposes in advance.

Second, traditional PIAs assume that data flows can be comprehensively mapped and documented. Yet AI systems, particularly those using federated learning or differential privacy techniques, may process data in ways that are inherently distributed and dynamic. The training process itself transforms data in ways that may not be fully understood even by the system's designers, while inference operations may reveal unexpected connections between seemingly unrelated data points.[9]

Third, the risk assessment methodology in traditional PIAs focuses primarily on direct harms from unauthorized access, use, or disclosure of personal data. This approach fails to capture the indirect but potentially more significant risks arising from algorithmic decision-making, such as discriminatory outcomes, manipulation through targeted content, or the creation of synthetic identities that can affect individuals who never consented to processing.

Finally, traditional PIAs treat privacy risks as static phenomena that can be identified and addressed through upfront design choices and ongoing monitoring. This assumption breaks down in the context of AI systems that continuously evolve through learning processes, where new risks may emerge from the interaction between the system and its environment rather than from predetermined processing operations.[10]

The GDPR's attempt to address algorithmic processing through Article 22's automated decision-making provisions and the requirement for "suitable safeguards" has proven similarly inadequate. The right to explanation, often cited as a key protection for algorithmic subjects, remains poorly defined and difficult to implement in practice, particularly for complex machine learning models that may not provide meaningful explanations even to their operators. The result is a framework that provides the appearance of protection while failing to address the fundamental challenges posed by AI systems to individual privacy and autonomy.[11]

[9] Gellert, R. "Data Protection Impact Assessments: A Meta-Regulatory Approach," Computer Law & Security Review, 2020.
[10] Casey, Farhangi, & Vogl. "Rethinking Explainability and Regulability in Algorithmic Governance," SSRN 2022.
[11] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling," October 2017.

III. The AI Challenge: Why Traditional PIAs Fall Short

The fundamental challenge facing traditional PIAs in AI contexts stems from the collision between regulatory frameworks designed for predictable, linear data processing and technologies characterized by complexity, opacity, and emergent behaviors. AI systems exhibit three core characteristics that render traditional assessment methodologies inadequate: dynamic processing behaviors, probabilistic outputs, and emergent properties that cannot be fully anticipated at the design stage.

Technical Inadequacies

The most immediate failure of traditional PIAs lies in their inability to capture the technical realities of modern AI systems. Unlike conventional data processing operations, AI systems, particularly machine learning models, do not simply transform inputs into predetermined outputs according to static rules. Instead, they engage in complex pattern recognition and generation processes that can produce unexpected results even when operating as designed.[12]

Generative AI systems exemplify this challenge. Large Language Models (LLMs) trained on vast datasets can produce outputs that combine information from multiple sources in ways that may inadvertently reveal sensitive information about individuals who contributed to the training data. Traditional PIAs, which focus on the direct use of identifiable personal data, fail to account for the risk of model inversion attacks or membership inference attacks that can extract information about training data subjects.

The problem is compounded by the black box nature of many AI systems. While traditional PIAs assume that processing operations can be described with sufficient detail to enable meaningful risk assessment, the internal workings of neural networks and other machine learning architectures often remain opaque even to their designers. This opacity makes it impossible to conduct the type of comprehensive risk analysis that traditional PIAs require.

Procedural Failures

Beyond technical limitations, traditional PIAs suffer from procedural failures when applied to AI systems. The linear, stage-gate approach that characterizes most PIA methodologies

[12] Selbst, A. & Powles, J. "Meaningful Information and Explanation in Automated Decision-Making," Harvard Journal of Law & Technology, 2017.

assumes that privacy risks can be identified and addressed during a discrete assessment phase, with periodic reviews sufficient to maintain protection over time. This approach fundamentally misunderstands the iterative nature of AI development and deployment.

AI systems typically undergo continuous training, fine-tuning, and updates that can significantly alter their behavior and risk profile. A language model that initially produces benign outputs may begin generating biased or harmful content after being exposed to new training data, while a recommendation system may develop discriminatory patterns as it learns from user interactions.[13] Traditional PIAs, conducted at fixed intervals, cannot capture these dynamic changes in risk.

The stakeholder engagement processes embedded in traditional PIAs also prove inadequate for AI systems. The GDPR requires consultation with data subjects and other stakeholders as part of the DPIA process, but AI systems often affect individuals and communities in ways that are not immediately apparent. The effects of algorithmic bias may not manifest until the system has been deployed at scale, while the indirect effects of AI systems on democratic discourse or social cohesion may only become visible over extended periods.[14]

Contextual Blindness

Perhaps most critically, traditional PIAs exhibit a form of contextual blindness that renders them particularly unsuited to AI governance. The focus on data protection compliance, while important, fails to capture the broader implications of AI systems for human rights and social justice. Algorithmic systems can violate principles of non-discrimination, fairness, and human dignity even when they technically comply with data protection requirements.[15]

This limitation is particularly evident in the treatment of automated decision-making under current frameworks. Article 22 of the GDPR provides limited protection against automated decisions with "legal or similarly significant effects," but its narrow scope excludes many AI applications that may have profound impacts on individuals' life opportunities. A hiring algorithm that systematically disadvantages certain demographic groups may not trigger

[13] Arvind Narayanan et al., "Membership Inference Attacks Against Machine Learning Models," IEEE S&P, 2018.
[14] Burrell, J. "How the machine 'thinks': Understanding opacity in machine learning algorithms," Big Data & Society, 2016.
[15] European Union Agency for Fundamental Rights, "Getting the future right: Artificial intelligence and fundamental rights," FRA, 2020.

Article 22's protections if the final hiring decision involves some human oversight, yet the discriminatory effects remain unchanged.[16]

The contextual blindness of traditional PIAs also manifests in their treatment of collective and societal harms. While privacy law has traditionally focused on individual rights and harms, AI systems often operate at population scale with effects that transcend individual privacy concerns. The manipulation of information ecosystems through algorithmic curation can undermine democratic deliberation, while the concentration of AI capabilities in a small number of powerful actors can create systemic risks to competition and innovation.[17]

Current PIA frameworks lack the conceptual tools and methodological approaches necessary to address these broader implications of AI deployment. The result is a form of tunnel vision that focuses narrowly on traditional privacy harms while ignoring the more significant risks that AI systems may pose to individual autonomy and social welfare.

IV. Empirical Evidence of PIA Failures

The theoretical limitations of traditional PIAs in AI contexts are increasingly supported by empirical evidence of their failure to prevent or adequately address significant privacy and rights violations. Analysis of recent AI incidents, enforcement actions, and regulatory responses reveals a consistent pattern: organizations conducting nominally compliant PIAs while deploying systems that cause substantial harm to individual rights and societal interests.[18]

High-Profile AI Incidents

The ChatGPT conversation leak incident of March 2023 provides a paradigmatic example of PIA failure in practice. OpenAI's system experienced a bug that allowed users to view snippets of conversations from other users' chat histories, potentially exposing sensitive personal information shared in confidence with the AI system. Despite OpenAI having conducted privacy assessments for ChatGPT, these assessments failed to anticipate or prevent a

[16] ICO, "What are the accountability and governance implications of AI?" 2024.
[17] Taylor, L., Floridi, L., & van der Sloot, B. "Group privacy: New challenges of data technologies," Springer, 2017.
[18] Commission Nationale de l'Informatique et des Libertés (CNIL), Carrying Out a Data Protection Impact Assessment if Necessary, at 4–5 (2025), https://www.cnil.fr/en/carrying-out-protection-impact-assessment-if-necessary.

vulnerability that arose from the complex interaction between the system's conversation management infrastructure and its user interface.[19]

More significantly, the incident revealed how traditional PIA methodologies fail to account for the emergent risks of AI systems operating at scale. The privacy breach did not result from a failure of the core AI model or a straightforward data handling error, but from the interaction between multiple system components in ways that could not have been fully anticipated through traditional risk assessment approaches. The incident affected millions of users worldwide and led to temporary suspension of the service, yet it fell outside the scope of conventional privacy risk categories.

Facial recognition bias incidents provide another category of evidence for PIA inadequacy. The Gender Shades study by Buolamwini and Gebru documented significant accuracy disparities in commercial facial recognition systems, with error rates of up to 34.7% for dark-skinned women compared to 0.8% for light-skinned men. Despite these systems having undergone various forms of assessment and validation, the discriminatory impacts were only discovered through external auditing rather than internal risk assessment processes.[20]

The systematic nature of these failures suggests that the problem extends beyond individual implementation errors to fundamental methodological limitations. Traditional PIAs focus on protecting data subjects from unauthorized access or disclosure, but they lack frameworks for assessing algorithmic fairness or detecting systemic bias that may emerge from seemingly neutral technical choices.

Enforcement Statistics and Regulatory Responses

Quantitative analysis of enforcement actions provides additional evidence of the enforcement gap. Despite the proliferation of AI systems across multiple sectors, privacy authorities have issued relatively few enforcement actions specifically addressing AI-related privacy violations. This enforcement deficit reflects not organizational compliance but rather the inadequacy of existing legal frameworks and assessment methodologies to capture AI-specific harms.

[19] James Vincent, OpenAI Fixes ChatGPT Bug That Leaked User Conversations, The Verge (Mar. 24, 2023), https://www.theverge.com/2023/3/24/23654814/openai-chatgpt-bug-user-conversations-leak.
[20] Joy Buolamwini & Timnit Gebru, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, Proc. Mach. Learn. Res. 81:1, 8 (2018).

The European Data Protection Board's guidance on AI and data protection acknowledges these limitations while maintaining adherence to existing PIA frameworks. The EDPB recognizes that AI systems present "specific risks" including automated discrimination, fictitious content generation about real persons, and attacks specific to AI systems such as model inversion. However, the Board's recommended approach remains anchored to traditional DPIA methodologies, with additional considerations rather than fundamental reconceptualization.[21]

National regulatory responses reveal similar patterns of recognition without adequate adaptation. The UK Information Commissioner's Office has published extensive guidance on AI and data protection, acknowledging that AI can involve "several processing operations that are themselves likely to result in a high risk," yet the ICO continues to rely on existing DPIA frameworks with marginal modifications. The result is guidance that identifies the problems without providing effective solutions.[22]

Case Studies in Sectoral Implementation

Healthcare AI deployment provides particularly compelling evidence of PIA inadequacy. AI systems used for medical diagnosis, treatment recommendation, and patient monitoring process highly sensitive personal data with potentially life-or-death consequences, yet traditional PIAs focus primarily on data security and access controls rather than algorithmic accuracy and bias. A recent case study of DPIA implementation for an AI-based healthcare software system revealed significant gaps between assessment requirements and actual privacy risks.

The healthcare case study demonstrated that traditional DPIA methodologies failed to address critical risks including algorithmic bias in diagnosis, inappropriate automation of clinical decision-making, and the potential for AI systems to perpetuate or exacerbate health disparities. While the DPIA process identified routine privacy risks such as data breach vulnerabilities, it provided no framework for assessing whether the AI system would produce equitable outcomes across different patient populations.

Financial services AI applications present similar challenges. AI systems used for credit

[21] Sandra Wachter, Brent Mittelstadt & Luciano Floridi, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation, 7 Int'l Data Privacy L. 76, 80 (2017).
[22] U.K. Information Commissioner's Office, Conducting a Data Protection Impact Assessment, at 4–5 (2021), https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpia/.

scoring, fraud detection, and algorithmic trading operate in highly regulated environments with extensive privacy protections, yet continue to produce discriminatory outcomes that fall outside the scope of traditional PIAs. According to recent analysis, 91% of financial institutions are actively using AI, yet only 28% have formal AI governance frameworks in place. This gap between AI adoption and governance maturity reflects the inadequacy of existing assessment methodologies to address AI-specific risks.

The Compliance Paradox

Perhaps most troubling is the emergence of what might be termed a compliance paradox: organizations can demonstrate technical compliance with PIA requirements while deploying AI systems that cause significant harm to individual rights and social welfare. This paradox reflects the fundamental mismatch between regulatory frameworks designed for traditional data processing and the realities of AI system deployment.

The paradox is evident in the widespread adoption of privacy-washing practices, where organizations conduct perfunctory PIAs that check regulatory boxes without meaningfully addressing AI-specific risks. These assessments often focus on data minimization and security measures while ignoring algorithmic bias, system transparency, and long-term societal impacts. The result is a false sense of security that may actually impede the development of more effective governance mechanisms.

V. Emerging Alternatives: AIAs and FRIAs

Recognition of traditional PIA limitations has spurred development of alternative assessment frameworks specifically designed for AI systems. Algorithmic Impact Assessments (AIAs) and Fundamental Rights Impact Assessments (FRIAs) represent the most significant attempts to bridge the governance gap, offering expanded scope and novel methodologies for AI risk assessment. However, analysis reveals that these emerging frameworks, while promising, face their own limitations and implementation challenges.[23]

Algorithmic Impact Assessments: Scope and Methodology

AIAs emerged from recognition that traditional privacy assessments fail to capture the broader

[23] Joshua A. Kroll et al., Accountable Algorithms, 165 U. Pa. L. Rev. 633, 646–53 (2017).

implications of algorithmic decision-making. Unlike PIAs, which focus primarily on data protection compliance, AIAs adopt a more holistic approach that considers algorithmic fairness, transparency, accountability, and societal impact. The methodology extends beyond privacy risks to encompass bias detection, performance evaluation across demographic groups, and assessment of system impacts on democratic values and social justice.[24]

The Canadian Algorithmic Impact Assessment tool exemplifies this expanded approach, utilizing 65 risk questions and 41 mitigation questions to evaluate automated decision systems across multiple dimensions. The tool assesses not only privacy implications but also algorithmic accuracy, fairness across different populations, and the potential for discriminatory outcomes. This multidimensional approach represents a significant advance over traditional PIAs in capturing AI-specific risks.[25]

However, empirical analysis reveals significant limitations in AIA implementation. A comparative study of algorithmic impact assessments and AI audits found that while AIAs provide valuable documentation of potential effects, they often lack the technical depth necessary to detect sophisticated forms of algorithmic bias or manipulation.[26] The assessments tend to rely on self-reporting by system developers rather than independent technical evaluation, creating opportunities for assessment gaming where organizations provide optimistic evaluations that may not reflect actual system performance.[27]

Fundamental Rights Impact Assessments: The EU AI Act Approach

The EU AI Act's introduction of FRIAs represents the most comprehensive attempt to date to develop AI-specific assessment methodologies. Article 27 requires deployers of high-risk AI systems to perform fundamental rights impact assessments that go beyond traditional privacy concerns to encompass the full range of rights protected under EU law. The FRIA framework explicitly recognizes that AI systems can affect rights to non-discrimination, freedom of

[24] Andrew D. Selbst, An Institutional View of Algorithmic Impact Assessments, 36 Harv. J.L. & Tech. 557, 563–67 (2023).
[25] Government of Canada, Algorithmic Impact Assessment Tool, https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html (last updated May 29, 2024).
[26] Derek E. Bambauer, Comparing Algorithmic Impact Assessments and AI Audits, SSRN, at 8–12 (June 19, 2023), https://ssrn.com/abstract_id=4486259.
[27] Law Commission of Ontario, Human Rights AI Impact Assessment, https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/human-rights-ai-impact-assessment/ (Mar. 2, 2025).

expression, human dignity, and democratic participation.[28]

The FRIA methodology incorporates several innovations over traditional PIAs. Rather than focusing solely on data protection, FRIAs require systematic evaluation of impacts on fundamental rights, including indirect and cumulative effects that may not be immediately apparent. The assessment must consider the specific categories of persons likely to be affected, the risks of harm taking into account information provided by AI system providers, and the implementation of human oversight measures.[29]

Recent research has developed sophisticated methodological frameworks for FRIA implementation, including the HH4AI Methodology which provides a structured, gate-based approach to assessing AI systems' human rights impacts. This framework employs filtering mechanisms that tailor assessments to specific system characteristics while maintaining comprehensive coverage of potential rights impacts. The methodology has been demonstrated through healthcare AI case studies, showing promise for systematic risk identification and mitigation planning.

Comparative Analysis: Strengths and Limitations

Comparative analysis of traditional PIAs, AIAs, and FRIAs reveals both significant advances and persistent limitations in emerging frameworks. AIAs and FRIAs demonstrate clear superiority in scope comprehensiveness, addressing algorithmic bias, fairness, and broader societal impacts that traditional PIAs ignore. The expanded stakeholder engagement requirements in these frameworks also represent an important advance, recognizing that AI systems affect communities and social groups rather than just individual data subjects.

However, several critical limitations remain across all current assessment frameworks. First, the problem of dynamic systems persists even in advanced frameworks. While FRIAs acknowledge that AI systems evolve over time, the assessment methodologies remain fundamentally static, conducted at discrete points rather than providing continuous monitoring of system behavior and impacts.

[28] Regulation (EU) 2022/123 of the European Parliament and of the Council, Artificial Intelligence Act, art. 27, 2023 O.J. (L 117) 1 [hereinafter EU AI Act].
[29] Gianclaudio Malgieri & Enrico Comandé, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76, 77 (2017).

Second, technical complexity barriers limit the effectiveness of all current approaches. The mathematical and computational sophistication required to understand modern AI systems often exceeds the expertise available to assessment teams, leading to superficial evaluations that miss critical technical vulnerabilities. This limitation is particularly acute for generative AI systems and large language models, where the relationship between training data, model architecture, and output behavior remains poorly understood even by technical experts.

Third, enforcement and accountability mechanisms remain underdeveloped across all frameworks. While AIAs and FRIAs provide more comprehensive risk identification than traditional PIAs, they offer limited guidance on how to address identified risks or hold organizations accountable for assessment quality. The result is often elaborate documentation exercises that provide limited practical protection for affected individuals and communities.

Institutional and Implementation Challenges

Analysis of early FRIA implementation reveals significant institutional capacity constraints that may limit the effectiveness of even well-designed assessment frameworks. The EU AI Act's FRIA requirements will apply to thousands of organizations across multiple sectors, yet regulatory authorities lack the technical expertise and resources necessary to evaluate assessment quality or verify compliance.[30]

The fragmentation of assessment approaches across different jurisdictions and sectors also poses challenges for organizations operating across multiple regulatory environments. While the EU emphasizes fundamental rights protection, the United States focuses on algorithmic accountability, and other jurisdictions develop their own approaches to AI governance. This regulatory fragmentation creates compliance complexity and may undermine the development of coherent global standards for AI assessment.

Perhaps most significantly, the pace of AI development continues to outstrip regulatory adaptation. The emergence of increasingly sophisticated generative AI systems, multimodal models, and autonomous AI agents presents new challenges that existing assessment frameworks including the most advanced AIAs and FRIAs, are not designed to address. The

[30] European Data Protection Supervisor, Guidelines on Artificial Intelligence and Data Protection, at 11 (2024), https://edps.europa.eu/data-protection/our-work/publications/guidelines/artificial-intelligence-ai_en.

result is a persistent gap between regulatory frameworks and technological reality that may require more fundamental reconceptualization of AI governance approaches.

VI. A Framework for Next-Generation AI Governance

The empirical evidence of traditional PIA failures and the limitations of emerging alternatives point toward the need for a more fundamental reconceptualization of AI governance frameworks. Rather than incremental improvements to existing assessment methodologies, effective AI governance requires a dynamic, adaptive, and outcome-focused approach that can evolve with AI systems themselves while maintaining meaningful protection for individual rights and societal interests.

Principles for Next-Generation AI Governance

Any effective AI governance framework must be grounded in principles that acknowledge the unique characteristics of AI systems while maintaining fidelity to fundamental values of human rights protection and democratic accountability. Four core principles should guide the development of next-generation approaches[31]:

Adaptive Governance: Unlike traditional regulatory approaches that rely on static rules and periodic assessments, AI governance must be inherently adaptive, capable of evolving as AI systems and their impacts change over time. This requires moving beyond point-in-time assessments toward continuous monitoring and evaluation mechanisms that can detect emerging risks and adapt governance responses accordingly.

Outcome-Focused Accountability: Rather than focusing primarily on procedural compliance with assessment requirements, next-generation frameworks should emphasize actual outcomes and impacts on individual rights and social welfare. This shift requires developing robust metrics for algorithmic fairness, transparency, and social benefit that can be measured and verified independently.

Multi-Stakeholder Participation: Effective AI governance cannot be achieved through top-down regulatory imposition alone but requires meaningful participation from affected communities, technical experts, civil society organizations, and other stakeholders. This

[31] Organisation for Economic Cooperation and Development (OECD), Recommendation of the Council on Artificial Intelligence, OECD Legal Instrument No. 0449 (May 22, 2019).

participatory approach is essential for capturing the full range of AI impacts and ensuring that governance mechanisms reflect diverse perspectives and values.

Technical Sophistication: AI governance frameworks must incorporate sufficient technical depth to address the actual mechanisms through which AI systems operate and cause harm. This requires developing governance institutions and processes that can engage meaningfully with technical complexity rather than treating it as a black box.

Institutional Architecture for Dynamic AI Governance

Implementing these principles requires new institutional architectures that can provide both technical expertise and democratic legitimacy in AI governance. The traditional model of regulatory agencies issuing static rules and conducting periodic compliance reviews is fundamentally inadequate for the dynamic nature of AI systems.

AI Governance Boards should be established as specialized institutions with the technical expertise and institutional capacity necessary for effective AI oversight. These boards should combine technical specialists, ethicists, legal experts, and community representatives in governance structures that can provide both scientific rigor and democratic accountability. The boards should have authority to require continuous monitoring, investigate incidents, and mandate changes to AI systems that pose unacceptable risks.[32]

Algorithmic Auditing Infrastructure represents another critical institutional innovation. Rather than relying on self-assessment by AI developers, effective governance requires independent technical evaluation of AI systems by qualified auditors. This infrastructure should include standardized auditing methodologies, certification programs for auditors, and institutional mechanisms for ensuring audit quality and independence.[33]

Participatory Assessment Mechanisms should be developed to ensure meaningful stakeholder engagement in AI governance. These mechanisms should go beyond traditional public comment processes to include deliberative forums, citizen panels, and community-based

[32] World Economic Forum, Global Technology Governance: AI in Action, at 38 (2025), https://reports.weforum.org/docs/WEF_AI_in_Action_Beyond_Experimentation_to_Transform_Industry_2025.pdf.

[33] Shreya Singh, AI Governance Beyond 2025: UN Pathways and Implications, Graduate Inst. Geneva, at 44–47 (Sept. 2025), https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2----Shreya-Singh.pdf.

monitoring programs that can provide ongoing input on AI system impacts. Special attention should be paid to ensuring participation from communities that are often marginalized in traditional governance processes but disproportionately affected by AI systems.

Methodological Innovations for AI Assessment

The methodological approach for next-generation AI governance should incorporate several key innovations that address the limitations of current assessment frameworks[34]:

Continuous Risk Monitoring: Rather than conducting assessments at discrete points in time, AI governance should implement continuous monitoring systems that can detect changes in AI system behavior and impacts in real-time. This approach should utilize automated monitoring tools, stakeholder feedback mechanisms, and regular performance audits to maintain ongoing awareness of system behavior.

Algorithmic Impact Metrics: Effective AI governance requires developing standardized metrics for measuring algorithmic fairness, transparency, and social impact that can be applied consistently across different systems and contexts. These metrics should be technically rigorous while remaining accessible to non-technical stakeholders.

Scenario-Based Assessment: Given the difficulty of predicting all possible impacts of AI systems, assessment methodologies should incorporate scenario-based approaches that evaluate system behavior under a range of potential conditions. This approach should include stress testing for edge cases, adversarial conditions, and long-term systemic effects.

Rights-Based Impact Analysis: Building on the FRIA approach, next-generation frameworks should incorporate comprehensive analysis of AI impacts on the full range of human rights, including civil, political, economic, social, and cultural rights. This analysis should consider both direct and indirect effects, as well as cumulative impacts across multiple AI systems.[35]

Implementation Roadmap

Transitioning from current governance approaches to next-generation frameworks requires a

[34] European Parliament, Policy Department for Economic, Scientific and Quality of Life Policies, The Impact of the General Data Protection Regulation (GDPR) on AI (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf.
[35] U.N. Special Rapporteur on the Right to Privacy, Report on Privacy in the Digital Age, U.N. Doc. A/HRC/49/31, at 12–15 (Mar. 2022).

carefully planned implementation strategy that acknowledges both the urgency of AI governance challenges and the practical constraints facing regulatory institutions.

Phase 1: Pilot Programs and Capacity Building (6-12 months): Regulatory authorities should initiate pilot programs testing next-generation governance approaches in selected high-risk AI applications. These pilots should focus on developing institutional capacity, testing assessment methodologies, and building stakeholder engagement mechanisms.

Phase 2: Framework Development and Standardization (12-24 months): Based on pilot program results, regulatory authorities should develop comprehensive frameworks for next-generation AI governance, including standardized assessment methodologies, institutional structures, and accountability mechanisms. This phase should include extensive stakeholder consultation and international coordination to ensure coherent approaches across jurisdictions.

Phase 3: Full Implementation and Adaptive Management (24+ months): The final phase should involve full implementation of next-generation governance frameworks with built-in mechanisms for continuous improvement and adaptation. This phase should include regular evaluation of framework effectiveness, adjustment of methodologies based on experience, and expansion to additional AI application domains.

VII. Conclusion

The enforcement gap between traditional Privacy Impact Assessments and the realities of AI system deployment represents more than a technical regulatory challenge, it reflects a fundamental misalignment between governance frameworks designed for a simpler technological era and the complex, dynamic, and often opaque systems that now shape critical decisions affecting millions of individuals worldwide. This Article's analysis demonstrates that neither incremental reforms to existing PIA methodologies nor the emerging alternatives of AIAs and FRIAs adequately address the core challenges posed by AI systems to individual rights and democratic governance.

The empirical evidence is clear: traditional PIAs systematically fail to capture AI-specific risks, prevent significant privacy and rights violations, or provide meaningful accountability for algorithmic harms. The 56.4% increase in AI incidents in 2024, coupled with declining public trust and fragmented regulatory responses, underscores the urgency of developing more

effective governance approaches. While emerging frameworks like FRIAs represent important advances in scope and methodology, they remain anchored to static assessment paradigms that cannot keep pace with the dynamic nature of AI systems.[36]

The path forward requires embracing the fundamental insight that AI governance must be as adaptive and sophisticated as the systems it seeks to govern. This means moving beyond the comfortable certainty of compliance checklists toward dynamic, outcome-focused approaches that prioritize actual protection of human rights over procedural regularity. It requires building institutional capacity for technical sophistication while maintaining democratic accountability and meaningful stakeholder participation.

The framework proposed in this Article, emphasizing adaptive governance, outcome-focused accountability, multi-stakeholder participation, and technical sophistication—provides a roadmap for this transition. However, the framework's implementation will require sustained commitment from regulatory authorities, civil society organizations, and the AI industry itself to move beyond the current paradigm of governance theater toward substantive protection of individual rights and societal interests.

Future research should focus on developing the technical methodologies and institutional mechanisms necessary to operationalize next-generation AI governance frameworks. This includes creating standardized metrics for algorithmic fairness and transparency, designing participatory governance processes that can engage meaningfully with technical complexity, and establishing international coordination mechanisms to prevent regulatory fragmentation.

The stakes of this transition extend far beyond technical compliance or even individual privacy protection. As AI systems become increasingly central to economic, social, and political life, the adequacy of our governance frameworks will determine whether these powerful technologies serve human flourishing or exacerbate existing inequalities and undermine democratic values. The enforcement gap documented in this Article represents both a warning and an opportunity, a chance to build governance frameworks worthy of the AI age before the costs of continued failure become irreversible.

[36] G20, G20 AI Principles, Digital Economy Task Force (2019), https://www.g20-insights.org/policy_briefs/g20-ai-principles/.