The Italian Law Journal

Vol. 11 No. 02 (2025)

EDIZIONI SCIENTIFICHE ITALIANE

www.theitalianlawjournal.it

THE ITALIAN LAW JOURNAL

An International Forum for the Critique of Italian Law

Vol. 11 – No. 02 (2025)

www.theitalianlawjournal.it

ISSN 2421-2156 January 2026 DOI 10.23815/2421-2156.ITALJ

Table of Contents

The Italian Law Journal
Vol. 11 – No. 02 (2025)

Essays

Short Symposium 'Revisiting the Commons: Law, Crisis and Social Struggle'

Hard Cases

Corporate and Financial Markets Law

*Insights and Analyses**

The Italian Law Journal is sponsored by
'Harvard Italian Law Association' (HILA)
'Società Italiana degli Studiosi del diritto civile' (SISDIC)
'Società Italiana per la ricerca nel diritto comparato' (SIRD)
'National Italian American Bar Association' (NIABA)

Essays

A Post-Normal Science Framework for Rethinking EU AI Governance

Claudia Amodio*

Abstract

Taking its vantage point from the Post-Normal Science (PNS) framework, this paper examines how the governance of artificial intelligence (AI) has become a test case for democratic societies confronting conditions where 'facts are uncertain, values in dispute, stakes high, and decisions urgent.'

It traces how PNS, originally developed for other science-policy domains marked by inadequate expert-driven policymaking and contested categories of risk and safety, offers analytical and normative resources for reimagining AI governance. The framework's core principles - embracing irreducible uncertainty and constituting extended peer communities - find contemporary expression in United Nations Educational, Scientific and Cultural Organization's (UNESCO) participatory assessment methodologies which treat AI systems as sociotechnical assemblages requiring ethical scrutiny rather than merely technical certification. By contrast, the European Union (EU) AI Act, despite its ambitious scope, embeds a troubling contradiction: it recognizes AI's dynamic and unpredictable character while operationalizing oversight through risk categories, conformity assessments, and industry-led standardization processes that assume knowability and control.

Now that the EU AI Act has entered into force, the decisive arena for responsible and democratic AI governance has shifted from legislative debate to the seemingly quiet, procedural machinery of implementation through standards. Standardization emerges not as a neutral technical exercise but as a political process determining whose expertise matters, which harms register as regulable and what remains invisible to oversight. Here, corporate influence threatens to calcify into epistemic capture, encoding industry priorities as objective technical requirements. The paper argues that critical scholarship must engage these seemingly procedural spaces as sites where the material and epistemic foundations of rights and freedoms are actively being constructed - and where knowledge itself becomes both a source and outcome of law-making. Only through co-regulatory processes that embrace the 'uncomfortable knowledge' and epistemic humility demanded by PNS can Europe realize its stated ambition to steer AI's trajectory through a distinctive and responsible model of governance.

I. Introduction: Law, Science, and Power in an Age of Ecological and Technological Disruption

Contemporary societies are increasingly governed under conditions of epistemic pluralism and systemic risk.[1] As traditional certainties about authority and knowledge

* Associate Professor of Comparative Law, Ferrara University. This research was carried out within the framework of the Progetto di Ricerca di Interesse Nazionale (PRIN), Project 2020XBFME2.

[1] Epistemic pluralism refers to the coexistence of multiple, often competing, ways of knowing

erode, legal and political decision-making has become deeply entangled in struggles over how risks are defined, who holds the authority to define them, and which forms of expertise are deemed legitimate.[2] In this landscape, legal normativity operates less through definitive rules that resolve disputes and more through ongoing processes that manage scientific contestation, enable coordination among multiple stakeholders, and maintain legitimacy amid persistent uncertainty.[3] These evolving configurations are far from neutral. Underlying power asymmetries – between private and public actors, competing epistemic communities, or the Global North and South – continuously shape whose uncertainties are acknowledged and whose voices influence the boundaries of acceptable risk.[4]

Such tensions become especially salient in domains where legal and policy frameworks confront complex, long-term challenges - climate change, ecological degradation, or rapid technological transformation. While these phenomena are deeply interlinked and mutually reinforcing, they share a defining feature, namely, they emerge from processes and dynamics whose scope and consequences remain only partially understood. As a result, the interaction of evidence, argument, and persuasion stands at the core of contemporary regulatory governance.[5]

Within this broader landscape, ecological and digital transformations offer a valuable lens on the reconfiguration of law, science, and power, each illustrating a distinct mode of responding to persistent ambiguity and managing the politics of long-term risk. In both domains, trans-boundary challenges and fundamental trade-offs between innovation, risk, and social well-being are continually renegotiated across fragmented sites of authority shaped by transnational private actors and

and producing knowledge, which can inform or challenge policy decisions. Systemic risk denotes risks that emerge from complex interdependencies within social, technological, or ecological systems, where localized failures can propagate widely. The combination of these conditions means that policy-making increasingly involves navigating uncertainty, contested knowledge, and unforeseen consequences. See S. Jasanoff, *The Fifth Branch: Science Advisers as Policymakers* (Cambridge, MA: Harvard University Press, 1990); U. Beck, *World Risk Society* (Cambridge: Polity Press, 1999); B. Wynne, 'Uncertainty and Environmental Learning: Reconceiving Science and Policy in the Preventive Paradigm' 2 *Global Environmental Change*, 111-127 (1992); A. Stirling, 'Pluralising Progress: From Integrative Transitions to Transformative Diversity' 1 *Environmental Innovation and Societal Transitions*, 82-88 (2011).

[2] The COVID-19 pandemic vividly illustrated these dynamics, as debates over epidemiological models became sites of contestation over expertise, authority, and legitimate knowledge. See T. Rhodes and K. Lancaster, 'Mathematical Models as Public Troubles in COVID-19 Infection Control: Following the Numbers' *Health Sociology Review*, 1-18 (2020).

[3] The European Union's AI Act exemplifies this shift, relying extensively on harmonized standards and conformity assessment to operationalize abstract regulatory requirements amid rapid technological change and contested expertise. See nos 1 and 2 of para IV below.

[4] As aptly observed by H. Straßheim and P. Kettunen, 'When does evidence-based policy turn into policy-based evidence? Configurations, contexts, and mechanisms' 10(2) *Evidence & Policy: A Journal of Research, Debate and Practice*, 259, 260 (2014): 'More often than not, policy-relevant facts are the result of an intensive and complex struggle for political and epistemic authority on both sides; science as well as policy.'

[5] G. Majone, *Evidence, Argument, and Persuasion in the Policy Process* (New Haven: Yale University Press, 1989).

economic pressures. Both deploy legitimating discourses – 'sustainable development,' 'green economy,' 'ethical AI' – yet the approaches to governing irreducible knowledge gaps and long-term risks diverge in striking ways.

The key distinction lies in how each domain conceptualizes situations in which outcomes and causal relationships remain unclear.

Ecological governance has developed normative frameworks - most notably the precautionary principle - that formally acknowledge uncertainty both as a trigger for action and as an inherent feature of complex systems. This orientation reflects a recognition of systemic interdependence and the limits of predictive control in socio-ecological contexts.[6] Despite the political and economic constraints that often delimit its scope, the precautionary ethos of ecological governance, frequently invoked in the name of protecting future generations, embodies decades of institutional learning and legal evolution within the European Union (EU).[7]

In stark contrast, digital governance characteristically follows technological innovation rather than directing it, privileging technocratic flexibility, market competitiveness, and regulatory adaptability over precautionary approaches or democratic oversight. Artificial intelligence (AI) epitomizes this dynamic. Celebrated as an engine of economic growth and geopolitical influence, it advances despite unresolved questions about its long-term societal impacts. Profound uncertainties and potential harms to fundamental rights surround AI's systemic consequences. These include algorithmic bias and discrimination, erosion of human agency and autonomy, restrictions on freedom of expression and access to information, threats to privacy and data protection, impediments to access to justice, labor displacement, societal dependence on opaque infrastructures, and emerging forms of systemic vulnerability.[8] Yet, AI development proceeds at revolutionary pace, with calls to slow or reconsider prevailing policy trajectories remaining largely unheeded.[9] Rather than confronting this uncertainty as constitutive of the technology itself,

[6] On the challenges of operationalizing the precautionary principle, including the limited utility of cost-benefit analysis when risks and benefits cannot be reliably quantified, see RECIPES Consortium, 'Guidance on the Application of the Precautionary Principle in the EU' (Final Report, December 2021), available at https://tinyurl.com/bdkdzty5 (last visited 31 January 2026). The acronym RECIPES stands for 'REconciling sCience, Innovation and Precaution through the Engagement of Stakeholders.'

[7] For a recent overview of EU environmental policy, see European Parliament, 'Environmental Policy: General Principles and Basic Framework', available at https://tinyurl.com/5x47ccpb (last visited 31 January 2026).

[8] The understanding of vulnerability as both universal and context-dependent - varying according to the subjects involved, their situational circumstances, and the relational dynamics at stake - has been developed by the Digital Vulnerability in European Private Law (DiVE) project: C. Crea and A. De Franceschi eds, *The New Shapes of Digital Vulnerability in European Private Law* (Baden-Baden: Nomos, 2024); C. Amodio and A. Diurni eds, *Human Vulnerability in Interaction with AI in European Private Law* (Cham: Springer, 2025); C. Crea and M. Infantino eds, *Remedies to Digital Vulnerability in European Private Law* (Cham: Springer, forthcoming 2026).

[9] C. Rudschies and I. Schneider, 'The Long and Winding Road to Bans for Artificial Intelligence: From Public Pressure and Regulatory Initiatives to the EU AI Act' 4 *Digital Society*, 57 (2025).

AI governance typically construes it as a transient challenge to be managed through technological refinement, data accumulation, or enhanced predictive capacity.

The uneven institutionalization of the precautionary principle within the EU exemplifies this asymmetry. The Court of Justice of the European Union has repeatedly affirmed it as a general principle of EU law, but its application remains highly variable. No *a priori* boundaries exist for determining which risks or technologies warrant precautionary approaches, yet the principle's reach reflects the Union's broader struggle to reconcile its foundational commitments with competing imperatives to foster economic development, maintain adaptability, and sustain competitiveness under systemic uncertainty.

This asymmetry is all the more consequential because both domains grapple with overlapping concerns with respect to public health, individual rights, socio-technical imaginaries, and distributive struggles, all of which appertain to how risks and benefits are allocated. Crucially, the most urgent challenges in science-driven policy domains are not simply questions of what is known, but fundamentally about whose perspectives and values are recognized and included in these deeply value-laden decisions.[10]

In ecological governance, the precautionary principle intersects with an openness to plural knowledge and broader participation, fostering deliberative processes that seek to include diverse voices, including those traditionally marginalized. This inclusion is not merely procedural but reflects an acknowledgment that complex uncertainties and systemic risks require more than technical expertise - they demand ethical, social, and political engagement.

By contrast, digital governance remains largely technocratic and dominated by specialized experts and powerful corporate actors, often sidelining broader societal concerns. This narrow framing risks marginalizing critical perspectives on equity, justice, and democratic accountability. The result is a governance approach that leaves long-term systemic risks inadequately addressed and raises profound questions about legitimacy and public trust.

As environmental disruptions escalate and digital technologies continue to entangle themselves into human lives, this paper turns to Post-Normal Science (PNS) as a framework for rethinking governance in contexts where 'facts are uncertain, values in dispute, stakes high, and decisions urgent.'[11] Originally developed for environmental challenges, PNS prescribes precautionary measures and extended peer communities to address systemic risks. Applied to AI governance, it demonstrates that the complex, multi-layered risks of AI demand ethical and regulatory responses that integrate multiple legitimate perspectives and acknowledge modes of reflexive, democratized deliberation that fundamentally transcend

[10] S. Funtowicz and J. Ravetz, 'Knowledge, Power, and Participation in the Post-Normal Age' 237 *Ecological Economics*, 5 (2025).

[11] This formulation, recently described as the 'widely cited mantra' of post-normal science, originates in S. Funtowicz and J. Ravetz, 'Science for the Post-Normal Age' 25(7) *Futures*, 739 (1993). See Eid, 'Knowledge' n 10 above, 8.

technocratic risk management and industry-led standards, rather than minimize long-term uncertain impacts.

The paper consists of five sections. Section 2 outlines PNS as both a descriptive and prescriptive approach to governance, tracing its evolution from environmental policy to other science-driven regulatory domains. Section 3 applies this framework to artificial intelligence and examines the United Nations Educational, Scientific and Cultural Organization's (UNESCO) approach to AI regulation as a practical manifestation of PNS principles. Section 4 turns to the EU AI Act, critically analyzing its risk-based architecture through the lens of PNS and assessing the epistemic and normative implications of its standardization process. The conclusion proposes transformative pathways for developing a more distinctively European and responsible approach to the governance of AI.

II. Post-Normal Science: From Environmental Policy to Technology Governance

1. The Emergence of Post-Normal Science

PNS emerged in the 1990s as a response to profound shifts in both scientific practice and the broader societal landscape of science-driven domains.[12] For a long time, science was seen as a provider of absolute truth, continuously expanding our knowledge and control over the world. This perspective led to a simplistic view of the relationship between science and policy, where scientific findings were expected to directly inform and guide political decision-making. However, a series of events progressively challenged this perspective, with the development and use of atomic weapons during World War II being the pivotal moment that dramatically transformed how scientific knowledge intersects with political power and decision-making. Subsequent events - from the environmental crises of the 1960s highlighted in Rachel Carson's *Silent Spring* (1962), to nuclear accidents like Three Mile Island (1979) and Chernobyl (1986), and the Bovine Spongiform Encephalopathy (BSE) ('mad cow') crisis in the United Kingdom (UK) (1986-2001) - have further exposed the profound moral and political implications of scientific research, continuing the trajectory initiated by the atomic era.[13]

In response to emerging challenges, risk analysts initially turned to computer modeling of hazardous situations, assigning probabilities to potential adverse events. This approach aimed to provide policymakers with precise numerical guidance for decision-making. However, when uncertainties were factored into these calculations, a paradoxical situation often emerged, highlighting the limitations of traditional

[12] S. Funtowicz and J. Ravetz, *Uncertainty and quality in science for policy* (Dordrecht: Kluwer Academic Publishers, 1990); Eid, 'Science' n 11 above, 739-755. The 1990 book lays the groundwork, but the 1993 Futures article is widely regarded as the foundational statement of PNS.

[13] For historical analysis of the transformation of science-policy relations from the atomic era through environmental and health crises, see S. Jasanoff, *The Fifth Branch* n 1 above, 2-3.

risk assessment methods. This paradox is often described as 'zero-infinity' risk, brilliantly capturing scenarios where the probability of an event is small (approaching zero), yet its potential damage is immeasurably large (approaching infinity).[14]

Governing in the face of zero-infinity dilemmas exposes a critical limitation of technocratic approaches. When historical precedent is insufficient and potential consequences are catastrophic - such as in technological, environmental, or health crises - probabilistic assessment becomes epistemically problematic rather than merely technically challenging. What begins as an exercise in scientific objectivity quickly dissolves into value-laden political decisions about acceptable uncertainty, tolerable consequences, and the distribution of potential harms across populations and generations.

In 1972, nuclear physicist Alvin Weinberg encapsulated questions 'which can be asked of science and yet which cannot be answered by science' in his concept of 'trans-science'.[15] He contends that while the scientific community can, and must, delineate the factual, measurable components of hard issues, the boundary between what science can address and what remains unresolved is difficult to draw.

In the following years, these insights gained momentum, as the understanding of science as a human and social enterprise became central to the work of scholars such as Jerome Ravetz and Bruno Latour.[16] The emergence of PNS, as conceptualized by Silvio Funtowicz and Jerome Ravetz, reflected this critical epistemological and methodological turn.[17] Although the full public maturation of PNS came about two decades after their seminal defining paper (1993), its intellectual roots lie in the transformative experiences of scientists and activist movements during the 1970s which advocated for more participatory, reflexive, and socially responsible approaches to science and technology, particularly in areas such as environmental degradation, nuclear energy, public health, and military research.

Central to Funtowicz and Ravetz's concept was the insistence that 'the uncertainties of natural systems and the relevance of human values'[18] be fully integrated and made transparent in decision-making processes. This move led to an important corollary, namely that in some policy debates involving scientific input, it is not only unrealistic but potentially irresponsible to assume that all relevant

[14] See T. Page, 'Keeping Score: An Actuarial Approach to Zero-Infinity Dilemmas', in G.T. Goodman and W.D. Rove eds, *Energy Risk Management* (New York: Academic Press, 1979), exploring how such dilemmas arise in nuclear and energy risk, stressing how lack of historical precedent makes probabilistic prediction difficult when potential losses are catastrophic. See also Royal Society, *Risk: Analysis, perception and management* (London: The Royal Society, 1992); Y.Y Haimes, 'On the complex definition of risk: A systems-based approach' 29(12) *Risk Analysis*, 1647-1654 (2009); S. Funtowicz and J. Ravetz, 'Three types of risk assessment and the emergence of post-normal science', in S. Krimsky and D. Golding eds, *Social theories of risk* (Westport, CT: Praeger 1992), 251-273.

[15] A.M. Weinberg, 'Science and trans-science' 10(2) *Minerva*, 209-222 (1972).

[16] J. Ravetz, *Scientific Knowledge and Its Social Problems* (Oxford: Oxford University Press, 1971); B. Latour, *Science in Action* (Cambridge, MA: Harvard University Press, 1987).

[17] A. Saltelli and S. Funtowicz, 'What is science's crisis really about?' 91 *Futures*, 5-11 (2017).

[18] J. Ravetz, 'What is post-normal science?' 31 *Futures*, 647, 653 (1999).

aspects can be easily measured, or that the quality of scientific information could be presumed without scrutiny. Funtowicz and Ravetz theoretically identify a specific category of policy-related scientific situations that require novel and diverse methodological approaches. This category is distinguished by four defining attributes, articulated in the aforementioned and widely cited phrase: 'facts are uncertain, values are in dispute, stakes are high, and decisions are urgent'.[19]

This concise formulation not only encapsulates the essence of the complex scenarios PNS seeks to address but also underscores the challenges of decision-making in contexts where traditional approaches fall short. Indeed, through the evolution of the PNS framework, it became progressively clear that it could function simultaneously as both an analytical description and a normative prescription.

Accordingly, the term 'post-normal' marks a departure from two prevailing notions of normality.[20] The first is Thomas Kuhn's (1962) conception of *normal science* as puzzle-solving within a stable paradigm, assuming a shared cognitive and methodological consensus. The second is *normal policy-making*, grounded in the belief that expert-driven, specialized knowledge provides a sufficient basis for rational decision-making.

Although these two forms of normality operate in distinct spheres, scholars have observed that 'where in normal science the key task in interfacing science and policy is to get the facts right, in post-normal science this is complemented with a new key task of exploring the relevance of deep uncertainty and ignorance that limit our ability to establish objective, reliable, and valid facts'.[21] Moreover, while both forms of normality predominantly rest on a shared technocratic rationality and on the assumption that scientific and technological progress ultimately aligns with social progress,[22] PNS voices challenge this faith. In doing so, they highlight the limits of claims to technological neutrality and raise concerns about the growing concentration and privatization of knowledge, often prioritizing market returns over the public good.[23] One of the PNS's central arguments is that technocratic rationality and solutionism can only persist 'by virtue of banishing uncomfortable knowledge and the creation of implausible socio-technical imaginaries'.[24]

Should we seek an example at the intersection of the digital and ecological

[19] S. Funtowicz and J. Ravetz, 'Science' n 11 above, 744.

[20] J. Ravetz, 'What is' n 18 above, 648; S. Funtowicz and J. Ravetz, 'Post-normal science' *Internet Encyclopaedia of Ecological Economics* (International Society of Ecological Economics, 2003), available at https://tinyurl.com/yc7xcyx8 (last visited 31 January 2026).

[21] J.P. van der Sluijs, 'Uncertainty and dissent in climate risk assessment: A post-normal perspective' 7(2) *Nature and Culture*, 174, 177 (2012).

[22] J. Giuntoli et al, *Exploring new visions for a sustainable bioeconomy* (2023), available at https://tinyurl.com/8e7xp4bb (last visited 31 January 2026).

[23] S. Funtowicz and J. Ravetz, 'Knowledge' n 10 above, 4.

[24] M. Giampietro and S.G.F. Bukkens, 'Knowledge claims in European Union energy policies: unknown knowns and uncomfortable awareness' 91 *Energy Research & Social Science*, 11 (2022), building on the seminal work of S. Rayner, 'Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses' 41 *Economy and Society*, 107-125 (2012).

transition discourses, the widespread framing of AI as a solution to climate and ecological crises serves as a telling illustration. It exemplifies a form of technocratic optimism that PNS fundamentally critiques, one that tends to obscure the complex entanglements between technology, society, and environment. Against this backdrop, PNS invites a more nuanced and contextual understanding of AI. Rather than a neutral or inherently progressive tool, PNS beholds AI as a ubiquitous shaping force embedded within and reproducing the political and economic systems of extraction, production, consumption, and waste that underpin the very crises it purports to solve.[25]

Anchored in a model of anticipatory governance that values critical and reflexive approaches, PNS exposes the untenable assumptions sustaining prevailing legitimating narratives, allowing for alternative frames that are sensitive to power relations, the diversity of actors, and the plurality of interests and norms. In doing so, it opens up the possibility of recognizing alternative knowledge claims,[26] fosters dialogue across disciplinary communities in areas of high societal relevance,[27] and ultimately helps to challenge what has been called the 'unthinking consensus'.[28]

2. Core Principles: Precaution and Extended Peer Communities

In some of their more recent works, Funtowicz and Ravetz reflect on how PNS has evolved and gained influence across diverse scientific and policy domains since their seminal 1993 paper. In these works, they emphasize a specific shift in the social processes of science-from the pursuit of *truth* to the pursuit of *quality*.[29] As they explain, 'quality is achieved, not by a delusory pursuit of certainty, but by the skilled management of its uncertainties, involving all who have a concern for the issue.'[30]

This emphasis on managing uncertainty and involving stakeholders finds a natural counterpart in the precautionary principle, which emerged around the same

[25] For an in-depth exploration of AI's promise and perils in the environmental field, see United Nations Environment Programme (UNEP), AI has an environmental problem. Here's what the world can do about that, available at https://tinyurl.com/eb7na2ux (last visited 31 January 2026. See also D. Greenwood and R.J. Hougham, 'Mitigation and adaptation: Critical perspectives toward digital technologies in place-conscious environmental education' 13(1) *Policy Futures in Education*, 97-116 (2015).

[26] A concrete illustration of such alternative knowledge claims can be found in Z. Kovacic et al, *The Circular Economy in Europe: Critical Perspectives on Policies and Imaginaries* (London: Routledge, 2023) and K.A. Waylen et al, 'Post-normal science in practice: Reflections from scientific experts working on the European agri-food policy nexus' 141 Environmental Science & Policy, 158-167 (2023).

[27] M. Di Fiore et al, 'The Challenge of Quantification: An Interdisciplinary Reading' 61 *Minerva*, 53-70 (2023).

[28] S.G.F. Bukkens et al, *The Nexus Times* (Bergen: Megaloceros Press, 2020), 159.

[29] S. Funtowicz and J. Ravetz, 'Post-normal' n 20 above, 4.

[30] S. Funtowicz and J. Ravetz, 'Post-Normal Science: How Does It Resonate With the World of Today?', in V. Šucha and M. Sienkiewicz eds, *Science for Policy Handbook* (Amsterdam: Elsevier, 2020), 14, 18.

time as PNS was gaining theoretical traction.[31] It is no coincidence that the 1992 Rio Earth Summit and the subsequent 1998 Wingspread Statement, which famously articulated the precautionary principle, coincided with the growing influence of PNS. Both were part of a broader intellectual response to increasing recognition of complexity and societal tensions surrounding science and policy in the late 20th century.[32]

The principle challenges traditional scientific and regulatory approaches by shifting the central question from 'Can we definitively prove this is harmful?' to 'How can we responsibly manage potential risks given our current understanding?'.[33] First, it shifts the burden of proof, requiring those advocating for an activity to demonstrate its safety rather than placing the onus on critics to prove potential harm. This reversal contrasts with the conventional approach, which typically assumes safety until evidence of harm is presented. Furthermore, the precautionary principle embraces uncertainty, enabling policymakers to take proactive measures when potential risks are scientifically plausible, even without complete empirical evidence. This acceptance of uncertainty represents a paradigmatic move from deterministic scientific models to a more adaptive, contextual approach to knowledge production and risk management. Finally, unlike traditional scientific approaches that strive for value-neutrality, the precautionary principle explicitly integrates ethical considerations and societal values into the decision-making process.

As it is well known, the principle was formally enshrined in EU law with its inclusion in the Maastricht Treaty as part of environmental policy (EC Treaty, Art 174(2), and now appears in TFEU Art 191(2). The European Commission's 2000 Communication on the precautionary principle further clarified its scope and application, embedding it as a cornerstone of EU environmental and health governance.

Its influence is visible across multiple governance domains. Examples include water policy directives, restrictive regulations on genetically modified organisms,

[31] E.A. Rosa and T. Dietz, 'Metatheoretical foundations for post-normal risk' 1(1) *Journal of Risk Research*, 15-44 (1998); J. Ravetz, 'The post-normal science of precaution' 36(3) *Futures*, 347-357 (2004).

[32] A. Stirling, 'Science, precaution, and the politics of technological risk: converging implications in evolutionary and social scientific perspectives' 1128(1) *Annals of the New York Academy of Sciences*, 95-110 (2008).

[33] 'Although falling short of prescriptive decision rules, the precautionary principle does suggest a range of more modest, open-ended, but nonetheless highly effective methodologies and general qualities, which offer ways to complement and improve on conventional risk assessment. As such, it is clear that – contrary to common assertions – the precautionary principle is of practical relevance as much to risk assessment as to risk management. Precaution does not automatically entail bans and phase-outs, but instead calls for deliberate and comprehensive attention to contending policy or technology pathways. Far from being in tension with science, precaution offers a way to be more measured and rational about uncertainty, ambiguity and ignorance': A. Stirling 'Risk, precaution and science: towards a more constructive policy debate' 8(4) *EMBO Reports*, 309-315 (2007). See also J. Peel, *The Precautionary Principle in Practice: Environmental Decision-Making and Scientific Uncertainty* (Sydney: Federation Press, 2005).

and ambitious climate initiatives such as under the European Green Deal, which pre-emptively impose emissions reductions despite uncertainty in future climate models. In public health, the precautionary principle has shaped several chemical regulations and measures addressing food safety. During the BSE crisis, for instance, EU institutions relied on this principle to prioritize public health over economic interests, embodying a cautious approach in the face of scientific uncertainty regarding food safety. More recently, the COVID-19 pandemic further demonstrated the principle's relevance, as rapid preventive measures - such as travel restrictions and vaccine approvals - were implemented amid evolving scientific knowledge.

Crucially, the precautionary principle not only calls for attention to complex, long-term impacts that standard risk assessment methods may fail to capture but also advocates for broad participation and the inclusion of perspectives that extend beyond traditional forms of specialized expertise. As noted earlier, the connection between managing uncertainties through precautionary measures and involving all relevant stakeholders is central to the PNS framework. It is precisely in this context that the concept of the extended peer community, another key pillar of PNS, assumes particular significance.[34]

The extended peer community represents a further departure from traditional notions of scientific 'normality,' reinforcing the shift from seeking absolute scientific 'truth' to prioritizing quality. Mirroring the 'deliberative turn' in democratic theory, PNS reframes uncertainty as a collaborative opportunity to cultivate more robust, contextually nuanced understanding. The call for a transformation that embodies the wider *democratization of expertise* has propelled PNS into the spotlight of contemporary debates on science advice and the governance of risk and regulation.[35] The key insight is that in complex, high-stakes situations with profound uncertainties, recognizing multiple legitimate sources of knowledge is essential for comprehensive problem-solving and informed decision-making.

One significant manifestation is the Aarhus Convention, formally known as the United Nations Economic Commission for Europe (UNECE) Convention on access to information, public participation in decision-making, and access to justice in environmental matters. This international agreement has established a foundation for ensuring that citizens and stakeholders have a voice in shaping policies that affect their lives and environments, and has prompted ongoing discussions on public participation and stakeholder engagement in EU environmental politics. A prominent example can also be seen in the European Green Deal and related initiatives, which strongly emphasize that citizens should be a driving force in the transition to a sustainable economy. This involves engaging various stakeholders

[34] S. Funtowicz and J. Ravetz, 'Post-Normal Science: How Does' n 30 above, 18, and J. Ravetz, 'The Post-Normal' n 31 above, 347-357.

[35] C. Carrozza, 'Democratizing Expertise and Environmental Governance: Different Approaches to the Politics of Science and their Relevance for Policy Analysis' 17(1) *Journal of Environmental Policy & Planning*, 108-126 (2015), and J.P. van der Sluijs, n 21 above, 174-195 and F. Grinnell, 'Rethink our approach to assessing risk' 522 *Nature*, 257 (2015).

- including local communities, NGOs, and industry representatives - in discussions about environmental policies, thereby creating an extended peer community that contributes to policy formulation and implementation. The European Climate Pact, which encourages local authorities and citizens to collaborate on climate action initiatives related to the European Green Deal, further reinforces the role of extended peer communities in environmental governance by fostering partnerships between governments and local knowledge and experiences. Participatory approaches such as citizens' assemblies and focus groups extend this ethos not merely by involving diverse populations in discussions on sustainability issues, but by actively gathering input and insight from them. They exemplify what has been described as 'technologies of humility',[36] meaning institutional practices designed to confront uncertainty and complexity through openness, reflection, and inclusivity. By bringing together different forms of expertise, such participatory mechanisms foster broader democratic legitimacy and enable stakeholders to evaluate policy proposals through the lens of lived experience and local context.

Recent reflections by the European Commission's Group of Chief Scientific Advisors capture this growing appreciation for inclusive, multi-perspective engagement. They call for drawing

> 'on all good science from all scientific disciplines and perspectives that could contribute to the issue at hand, including the natural sciences, engineering, medicine, social sciences, and the humanities'.[37]

This signals a broader shift towards transdisciplinary and participatory forms of governance, resonating strongly with PNS's call to manage uncertainty collectively, through dialogue among all those who have a stake in the issue.

As the following section explores, this model of inclusive engagement provides a crucial and valuable template for AI governance.

III. Framing AI as a Post-Normal Challenge

1. The Unique Characteristics of AI Development

AI development is advancing at a pace that outstrips the ability of existing policy frameworks to keep up. It challenges conventional systems of risk management and control, including those introduced by the EU AI Act, which categorizes AI systems into risk levels but struggles to address systems that rapidly shift between

[36] S. Jasanoff, 'Technologies of Humility: Citizen Participation in Governing Science' 41(3) *Minerva*, 240 (2003), available at https://tinyurl.com/4zpekkrk (last visited 31 January 2026).

[37] European Commission: 'Group of Chief Scientific Advisors and Directorate-General for Research and Innovation, Scientific advice to European policy in a complex world' (Publications Office of the European Union, 2019) available at https://tinyurl.com/ynb7wezw (last visited 31 January 2026).

categories as they evolve or when deployed in different contexts.[38]

Much like steering a high-speed vehicle across uncharted terrain, AI's rapid development presents the fundamental challenge of harnessing its momentum while simultaneously anticipating what lies ahead, which is needed to provide direction and manage risks. As Floridi compellingly states, 'the best way to catch the technology train is not to chase it, but to be at the next station.'[39] Yet, the governance challenge is arguably compounded by dual layers of uncertainty.

First, there is fundamental uncertainty about the trajectory and implications of AI development itself. These are captured by the aforementioned 'zero-infinity' paradox where immediate harms or failures may appear negligible, yet the future consequences could be vast and unpredictable. AI's multilayered and dynamic nature has few historical parallels in scale, complexity, or societal reach. Rooted in science and engineering, yet extending far beyond them, AI transforms social, economic, and epistemic structures in ways that defy linear prediction. Even with growing technical sophistication, our understanding of how AI systems learn, generalize, and interact with their environments remains fragmentary - a reminder that much of AI's behavior is emergent rather than engineered.[40]

Second, there is policy uncertainty arising from the need to regulate amid incomplete and evolving knowledge of AI's capabilities and risks. While regulatory frameworks are emerging globally, the lessons learnt from social media's first two decades - algorithmic amplification of misinformation, surveillance capitalism, and erosion of epistemic commons - demonstrate how governance lag can allow technologies to reshape societal structures before adequate safeguards are established. With AI, this challenge intensifies, as policymakers must navigate a moving target where every decision carries systemic consequences: too much restraint risks stifling innovation and public value, while too little invites harm, concentration of power, and social disruption.

These underlying tensions exemplify the challenges anticipated by PNS in science-driven policy domains, where traditional governance approaches prove insufficient in contexts characterized by high stakes and deep uncertainty. AI's unique characteristics, particularly its technological exceptionalism and transformative economic potential, further amplify these governance complexities, rendering

[38] See nos 1 and 2 of para IV below.

[39] L. Floridi 'Soft Ethics and the Governance of the Digital' 31 *Philosophy & Technology*, 6 (2018).

[40] Addressing these issues has spurred the recent Explainable AI (XAI) movement, which seeks to make AI systems more transparent and interpretable to users. However, XAI approaches often focus on technical explanations rather than meaningful user understanding, highlighting the persistent challenge of bridging the human-AI divide. See L. Longo et al, 'A Manifesto of Open Challenges and Interdisciplinary Research Directions' 106 *Information Fusion*, 1-22 (2024); T. Miller, 'Explanation in Artificial Intelligence: Insights from the Social Sciences' 267 *Artificial Intelligence*, 1-38 (2019); T. Wang et al, 'From Aleatoric to Epistemic: Exploring Uncertainty Quantification Techniques in Artificial Intelligence' arXiv preprint arXiv:2501.03282 (2025), available at https://tinyurl.com/4pck4n7s (last visited 31 January 2026).

the stakes of both development and regulation exceptionally high.

The challenge is compounded by significant geopolitical dimensions. AI competition among the United States (US), China, and the EU extends far beyond technological innovation to encompass control over data infrastructure, semiconductor supply chains, computing power, and regulatory paradigms. China's strategic deployment of models like DeepSeek illustrates how AI development is increasingly intertwined with broader soft power objectives,[41] while recent multilateral efforts reveal deep fractures in global governance approaches. The February 2025 AI Action Summit in Paris exemplifies these tensions. While the summit's declaration emphasized that AI must be open, inclusive, transparent, ethical, safe, secure, and trustworthy, and stressed sustainability for both people and planet, the US and UK notably declined to sign. US Vice President J.D. Vance's criticism of Europe's regulatory approach as overly restrictive and his call for more 'innovation-friendly' governance underscores a fundamental divergence in how major powers balance innovation imperatives against precautionary principles.[42] These geopolitical fault lines suggest that AI governance will unfold not through unified global frameworks, but through competing regulatory models that reflect divergent values, risk tolerances, and strategic priorities.

2. AI and the Post-Normal Framework

The PNS fourfold characterization of *post-normal* situations - where 'facts are uncertain, values in dispute, stakes high, and decisions urgent' - captures with striking accuracy the conditions surrounding contemporary AI development. Viewing AI through this post-normal lens serves a dual purpose. First, it underlines the need for sustained global collaboration and transdisciplinary inquiry, aimed at cultivating a deeper, more qualitative understanding of both AI's internal mechanisms and its societal implications.[43] Second, it compels a reconfiguration of regulatory thinking, moving beyond ostensibly objective, evidence-based solutions toward an explicit recognition of the limits of certainty and the value-laden nature of technological governance.[44]

[41] O.J. Daniels and H. Dohmen, 'China's overlooked AI strategy: Beijing is using soft power to gain global dominance' Foreign Affairs, 25 July 2025, available at https://tinyurl.com/2j8m32hc (last visited 31 January 2026)

[42] 'UK and US refuse to sign international AI declaration', *BBC News*, 11 February 2025, available at https://tinyurl.com/kpfjr6cj (last visited 31 January 2026).

[43] E.M. Bender et al, 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2021), 613.

[44] See C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016), who goes so far as to propose that AI developers should adopt an ethical framework akin to the Hippocratic Oath taken by medical professionals, with a commitment to 'do no harm'. This framework would guide developers in making ethical decisions and foster transparency, continuous learning, and interdisciplinary collaboration, ensuring that AI serves the broader interests of society.

To be sure, applying the PNS model to the field of AI presents both opportunities and challenges. The strong competitive pressure both AI companies and governments face often leads to a reluctance to acknowledging uncertainties or potential limitations. This nurtures a narrative of technological inevitability and progress that tends to sideline the crucial goal of safeguarding public interests.[45] Moreover, major tech corporations and research institutions that dominate AI development are well known for prioritizing proprietary knowledge, often at the expense of fostering inclusive, multi-stakeholder approaches to knowledge production.[46] This concentration of power in the AI sector exacerbates societal imbalances and raises ethical concerns, while at the same time undermining the two core premises of PNS: adopting a precautionary regulatory approach and broadening scientific expertise beyond traditional technocratic boundaries.

Yet, despite these challenges, the dynamics already emerging in the evolving landscape of human-machine interaction underscore the critical need for inclusive and adaptive governance frameworks, namely ones capable of mitigating risks while fostering trust, accountability, and public legitimacy in the development and deployment of transformative AI technologies. As international discussions continue to unfold, a post-normal approach appears increasingly well-suited to this task, as it entails the delicate and unprecedented art of engaging with plural forms of knowledge and bridging scientific, ethical, and ideological divides. By creating a framework in which multiple understandings of AI's role in society can be openly considered, each reflecting the narratives we construct about ourselves, our relationship with nature, and our connections to others, can help ensure that AI ultimately serves broader, collective human interests on a truly global scale.

After all, most people engage with AI not as engineers or data scientists, but as social beings whose lives are shaped by its outputs and consequences. As observed,

> 'individuals are not concerned with the technical aspects of how an AI model is generated; rather, they are interested in understanding whether the model's results have implications for their own lives.'[47]

From this perspective, sustainable and ethically sound AI development requires much more than the continual expansion of datasets or computational power. Even a perfect understanding of an algorithm's technical design would not reveal the full range of its behaviors, effects, and unintended consequences, which only become apparent when these systems are embedded within complex social contexts and human relationships. As Ananny and Crawford remind us, understanding

[45] See generally S. Jasanoff, *The Ethics of Invention: Technology and the Human Future* (New York: W.W. Norton & Company, 2016).

[46] Y. Bengio et al, 'International AI Safety Report' arXiv preprint arXiv:2501.17805 (2025), 22.

[47] A. Bao and Y. Zeng, 'Understanding the Dilemma of Explainable Artificial Intelligence: A Proposal for a Ritual Dialog Framework' 11 *Humanities and Social Sciences Communications*, 321 (2024), available at https://tinyurl.com/4mtudwbh (last visited 31 January 2026).

algorithms entails more than simply peering inside their code.[48] The growing call for algorithmic transparency, while necessary, remains insufficient on its own to address the deeper ethical and epistemic challenges posed by AI systems.

A post-normal approach urges decision-makers to recognize that many risks associated with AI-such as bias, surveillance, or social manipulation-cannot be fully captured through quantifiable metrics alone. This broader engagement transforms AI governance into a collective learning process, one that navigates the complex interplay between technological innovation and human values, rather than treating them as separate or sequential domains.[49] Acknowledging the limits of certainty in AI research and development is especially crucial in light of emerging regulatory frameworks, such as the EU's AI Act. The processes of assessing risks, determining compliance, and defining accountability open new spaces for scientists, engineers, ethicists, and social scientists to assume advisory, auditing, and policy-shaping roles.[50] Stepping beyond the technical numbers and statistical machinery of a model, as well as critically examining the assumptions, interests, and values it embodies, does not make governance effortless. Nor does it guarantee that AI will naturally reflect the broader social, ethical, or cultural landscapes it touches. Nevertheless, adopting a reflective stance that resists 'quantification at all costs' is vital for cultivating AI governance that is alert, adaptive, and socially attuned - qualities that lie at the very core of the PNS framework.

Before turning to a critical analysis of the EU AI Act, the following section argues that UNESCO's endorsement of PNS principles in the AI domain lends further legitimacy to this perspective, underscoring the relevance of PNS in confronting the complex ethical and epistemic challenges of contemporary AI governance.

3. Implementing Post-Normal Principles in Global AI Governance: Insights from UNESCO's approach

While international debate on science-policy interfaces is increasingly engaging with PNS governance frameworks,[51] the United Nations' (UN) approach to AI, as reflected in its various initiatives and documents, is particularly relevant as it

[48] M. Ananny and K. Crawford, 'Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability' 20 *New Media & Society*, 973-989 (2018).

[49] T. Miller, 'Explanation in Artificial Intelligence: Insights from the Social Sciences' 267 *Artificial Intelligence*, 1-38 (2019).

[50] Andrea Saltelli, one of Italy's most prominent scholars in the field of science for policy - and author of the entry 'Scienza post-normale' for the Enciclopedia Treccani (2024) - has, together with Arnald Puy and Samuele Lo Piano, advanced the concept of sensitivity auditing, an extension of sensitivity analysis: 'Sensitivity auditing', in E. Padilla Rosa and J. Ramos-Martin eds, *Elgar Encyclopedia of Ecological Economics* (Cheltenham: Edward Elgar Publishing, 2023), 467.

[51] Besides the academic references already cited, see, among Italian scholars: A. Tavernaro, 'La science policy mancante: il COVID-19 in Italia e il dialogo tra istituzioni scienziati e cittadini' 46(6) *Epidemiologia & Prevenzione*, 456 (2021); M. Tallacchini, 'Scienza e potere', in M. Cartabia and M. Ruotolo eds, *Enciclopedia del Diritto, Potere e Costituzione* (Milan: Giuffrè Editore, 2023), 1059; A. Saltelli, 'Scienza post-normale' *Enciclopedia Treccani* (Roma: Treccani, 2024).

offers a compelling framework for conceptualizing AI as a post-normal challenge. More specifically, examining UNESCO's approach and its alignment with PNS principles helps illustrate how PNS concepts, originally developed in other policy domains, can inform a more comprehensive and inclusive framework for AI governance. Two key examples demonstrate this resonance: the Readiness Assessment Methodology (RAM)[52] and the Ethical Impact Assessment (EIA).[53]

a) The Readiness Assessment Methodology (RAM)

UNESCO's RAM closely reflects the PNS concept of the 'extended peer community,' particularly through its emphasis on inclusivity, participatory approaches, collaborative knowledge sharing, and multidimensional analysis. Currently piloted in the first cohort of 50 countries worldwide, RAM is an integral component of UNESCO's Recommendation on AI Ethics - the first global standard on AI ethics, developed through an inclusive, multistakeholder process in 2021 and applicable to all 194 member states of UNESCO.[54]

In its normative effort to help countries evaluate their AI ecosystem across legal, social, economic, scientific, and technological dimensions, RAM prioritizes inclusivity by establishing steering committees that bring together representatives from government, the private sector, academia, and civil society, resonating with the PNS call for involving stakeholders beyond traditional experts. Participatory tools such as questionnaires, feedback mechanisms, and interdisciplinary workshops are also integral to the RAM, reflecting PNS's emphasis on incorporating diverse knowledge and perspectives.

Moreover, the multidimensional assessment of AI ecosystems champions a holistic approach that fosters dialogue across sectors, echoing PNS's commitment to the recognition of systemic complexity. This comprehensive approach enables a thorough evaluation of national AI ecosystems, while the subsequent publication of results on UNESCO's Observatory and discussions at the Global Forum on the Ethics of AI serve as platforms for international knowledge exchange, allowing countries to share experiences and best practices during a period of rapid change in the AI governance landscape.[55]

This transparency plays a crucial role in other important multi-stakeholder

[52] UNESCO, 'Readiness assessment methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence' (Paris: UNESCO, 2023), available at https://tinyurl.com/2a9xwpx2 (last visited 31 January 2026)

[53] UNESCO, 'Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence' (Paris: United Nations Educational, Scientific and Cultural Organization, 2023), available at https://tinyurl.com/5x9k5vau (last visited 31 January 2026).

[54] UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (Paris: United Nations Educational, Scientific and Cultural Organization, 2021), available at https://tinyurl.com/4wenkyaz (last visited 31 January 2026).

[55] 'Global Forum on the Ethics of AI' (Bangkok: UNESCO, 24-27 June 2025 3rd edition), available at https://tinyurl.com/5cw4nfbp (last visited 31 January 2026).

platforms implemented within the UN's framework, such as the AI for Good platform,[56] the Commission on Science and Technology for Development,[57] and the Multistakeholder Forum on Science, Technology and Innovation for the Sustainable Development Goals.[58] The RAM is thus not just an assessment tool, but part of a broader, collaborative and interactive approach to understanding and managing AI's complex global implications.

As underscored by the UN White Paper on AI Governance, a document recently published that reflects deep concerns about the current technological governance landscape, the RAM's significance lies not merely in diagnostic capabilities, but in its potential to rebalance power dynamics increasingly dominated by the 'disproportionately large role of the private sector.'[59] The document emphasizes that by prioritizing multi-stakeholder collaboration and engaging marginalized communities, including those from the Global South, the UN aims to ensure that AI governance reflects a wide range of societal needs and values.[60]

b) The Ethical Impact Assessment (EIA)

Alongside the plea for inclusivity, collaborative learning and co-production of knowledge, the commitment to anticipatory mechanisms and 'future-proof' regulation is another core issue where UNESCO's Recommendation on the Ethics of AI and its implementation tools closely align with PNS. The above-mentioned White Paper provides compelling evidence of this alignment. It acknowledges

> 'the added value of risk and opportunity assessments that also duly incorporate human rights considerations and the need to adopt the precautionary principle when developing new technologies and AI.'[61]

The case for the adoption of the precautionary principle emerges most notably from the EIA and its methodology's emphasis on AI as a complex socio-technical system with profound and potentially unpredictable impacts. Grounded in the

[56] 'AI for Good' (Global Summit) (United Nations - International Telecommunication Union, 7-10 July 2026, Geneva) available at https://tinyurl.com/5n737z6a (last visited 31 January 2026).

[57] 'United Nations Commission on Science and Technology for Development (CSTD), Commission on Science and Technology for Development' (served by United Nations Conference on Trade and Development (UNCTAD)), available at https://tinyurl.com/3j956v3z (last visited 31 January 2026).

[58] '10th Multi-Stakeholder Forum on Science, Technology and Innovation for the Sustainable Development Goals' (STI Forum 2025) (United Nations US, 7-8 May 2025), available at https://tinyurl.com/5n7zrsmw (last visited 31 January 2026).

[59] 'United Nations System, United Nations System White Paper on AI Governance: An Analysis of Current Institutional Models and Related Functions and Existing International Normative Frameworks Within the United Nations System That Are Applicable to Artificial Intelligence Governance' (Geneva: United Nations Chief Executives Board for Coordination, 2024), 38, available at https://tinyurl.com/3kwps4sw (last visited 31 January 2026).

[60] ibid 23 and 42.

[61] ibid 38.

UNESCO Recommendation's values, the EIA fundamentally recognizes that AI systems do not exist in isolation, but are deeply entangled with 'intergenerational and historical biases in data collection relating to gender, race, culture and other factors.'[62]

In practice, this means systematically scrutinizing AI systems across several critical dimensions, so as to proactively identify both potential ethical risks and unintended societal consequences. To ensure a comprehensive assessment, this exercise is undertaken in consultation with a variety of stakeholders, including experts, developers, impacted and potentially impacted parties, etc.[63] This might involve, for example, analysing how a recommendation algorithm could potentially reinforce echo chambers, amplify misinformation, or create unexpected social dynamics. For instance, if a hiring algorithm shows disproportionate screening out of candidates from certain racial or gender groups, the EIA would flag these systemic biases.[64] In the same vein, an autonomous vehicle's assessment would go beyond technical performance to examine broader implications, from its interaction with existing transportation systems to its impact on employment and its role in ethical decision-making scenarios.[65]

Rooted in a 'model of anticipatory governance in which multidisciplinary collaboration, innovative thinking and transformational future-making are valued',[66] the EIA significantly advances traditional ethical oversight, promoting a culture of ethical responsibility that extends beyond binary assessments (pass/fail) and regulatory compliance and fosters dialogue and cooperation across cultural and institutional boundaries.[67] While it shares the anticipatory nature of conventional *ex ante* risk assessments, the EIA goes further by offering actionable recommendations to proactively address ethical concerns. Unlike traditional compliance mechanisms that deliver static judgments, the EIA thus operates as an ongoing process, integrating ethical considerations as dynamic and evolving components of governance and innovation.

Taken together, these developments not only reflect the growing acknowledgment of AI's post-normal nature but also signal a broader shift toward adaptive, quality-focused ethical governance, balancing forward-looking strategic vision with rigorous contextual grounding. In short, UNESCO's engagement with key principles of the PNS framework - particularly the precautionary principle and the involvement of extended peer communities - demonstrates the continued relevance of these ideas for confronting global AI challenges.

[62] 'United Nations System' n 59 above, 48.

[63] UNESCO, 'Ethical Impact Assessment' n 53 above, 45.

[64] ibid 23.

[65] ibid 25.

[66] 'United Nations System' n 59 above, 44.

[67] E. Bean et al, 'Eavesdropping on UNESCO AI Policy, Leadership, and Ethics' 18(4) *Journal of Leadership Studies*, 99-111 (2025).

IV. A Post-Normal Critique of the EU AI Act

1. Promises and Limitations of the EU AI Act Risk-Based Model

The historic enactment of the EU AI Act on 1 August 2024 represented a watershed moment in global technology regulation, culminating a marathon legislative process that sought to navigate the intricate and rapidly evolving AI landscape. As the first comprehensive legal framework for AI governance, the regulation has unsurprisingly become a focal point for debates on how AI should be regulated, the risks it poses, and how to balance innovation with ethics.

While the implementation phase only started on 2 February 2025, the EU AI Act has sparked contrasting viewpoints among academics, with critics falling into two main camps. On one side, there are those who fear the regulation prioritizes pre-emptive risk avoidance over flexibility and adaptability, potentially placing undue constraints on innovation and competitiveness;[68] on the other, those who contend that the proposed rules are insufficiently stringent and fail to adequately address potential risks associated with AI technologies.[69]

This debate resonates with broader theoretical and practical discussions in technology governance, particularly within the European context, where the precautionary principle has long informed regulatory approaches to emerging and uncertain risks. In 2020, the European Parliament explicitly advocated for the application of precaution to AI-based technologies as a reaffirmation of Europe's constitutional commitment to safeguarding human dignity, sometimes even above individual liberty.[70]

Yet a critical question remains: Does the EU AI Act fully embody the spirit and rigor of that 2020 call for precaution? A closer examination reveals a more nuanced reality.

On one hand, the Act's only explicitly precautionary measure is the ban on AI systems deemed to pose unacceptable risks.[71] On the other hand, its broader framework is largely anchored in conventional risk assessment methodologies, categorizing AI systems into risk levels and emphasizing quantifiable metrics and cost-benefit analyses. In effect, the Act treats AI through a dual lens: as dynamic agents whose societal impacts and implications for fundamental rights demand

[68] M. Draghi, *The Future of European Competitiveness: A Competitiveness Strategy for Europe* (Part A) (Brussels: European Commission, 2024), available at https://tinyurl.com/3zdrar6a (last visited 31 January 2026).

[69] G. Smorto, 'Human Vulnerability and AI. The Risk of Risk Regulation', in C. Amodio and A. Diurni eds, *Human Vulnerability* n 8 above.

[70] European Parliament, 'Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL))', 20 October 2020. See also G. De Gregorio, 'The Rise of Digital Constitutionalism in the European Union' 19(1) *International Journal of Constitutional Law*, 41-70 (2020).

[71] C. Rudschies and I. Schneider, 'The Long and Winding Road to Bans for Artificial Intelligence: From Public Pressure and Regulatory Initiatives to the EU AI Act' 4 *Digital Society*, 57 (2025).

vigilance, but also as certifiable objects that can be deemed safe before deployment.[72] Arguably the most problematic dimension of the Act is the creation of an exceptional legal framework for AI systems deployed by law enforcement, migration, and national security authorities. This framework is exceptional in that it allows these authorities to benefit from exemptions to the strict regulatory 'redlines' that apply to other high-risk AI systems, raising significant concerns about accountability, oversight, and the protection of fundamental rights.[73]

The EU AI Act establishes four risk levels - unacceptable, high, limited, and minimal - aiming to calibrate regulatory oversight to the potential impact of different AI applications. The Act sets out a clear yet relatively abstract list of typical high-risk applications in Annex III. Under this framework, the risk management system required by Art 9 applies only to systems explicitly designated as high-risk, and not to those that may become high-risk through misuse or overreach in particularly sensitive contexts. Consequently, the risk management obligations imposed on providers focus primarily on risk mitigation, rather than on the identification or assessment of emerging or unforeseen risks.[74]

The Act requires high-risk AI systems to undergo conformity assessments and continuous monitoring throughout their whole lifecycle, indicating a recognition of the provisional nature of AI safety and reliability. Yet, rather than questioning the fundamental appropriateness of deploying potentially harmful technologies, the Act treats these risks as technical problems that can be mitigated to 'acceptable' levels through a set of procedural safeguards. To put it bluntly, it imposes an ethical perspective where the label 'ethical' is little more than the new 'green', failing to fully engage with the uncertainties and long-term impacts of emerging AI technologies.[75]

Ultimately, the Act applies rules designed for static objects to technologies in constant flux. Risks, namely, the 'known knowns', can be assigned statistical probabilities and quantified effects. Uncertainties, by contrast, constitute 'known unknowns', that is, we recognize potential effects but cannot quantify them. Beyond these lie 'unknown unknowns', where we remain unaware that certain technologies or activities may cause harm at all. Compounding these challenges, the Act's emphasis on minimizing known hazards relies heavily on technical standards whose interpretation and enforcement is largely delegated to those directly involved

[72] M. Almada and N. Petit, 'The EU AI Act: A Medley of Product Safety and Fundamental Rights?', in *EUI RSC Working Paper 2023/59* (Florence: European University Institute, 2023), available at https://tinyurl.com/ybx4rxn8 (last visited 31 January 2026).

[73] ibid 13.

[74] M. Ebers, 'Truly Risk-based Regulation of Artificial Intelligence How to Implement the EU's AI Act' 16(2) *European Journal of Risk Regulation*, 684, 693 (2025). This approach is embedded in several provisions throughout the regulation such as Recital 30, Art 4, Art 9, Art 13.

[75] M.D. Schultz et al, 'Digital Ethicswashing: A Systematic Review and a Process-Perception-Outcome Framework' 5 *AI and Ethics*, 805-818 (2025); V.C. Müller, 'Ethics of Artificial Intelligence and Robotics', in E.N. Zalta ed, *Stanford Encyclopedia of Philosophy* (Palo Alto: CSLI, Stanford University, 2020), 1-70.

in AI development and deployment. Such an approach potentially leaves critical gaps in addressing the unforeseen risks associated with high-risk AI systems.[76]

In this context, the Act's introduction of regulatory sandboxes (Arts 57, 58, and 59) represents an experimental solution in a world where technological developments outpace regulation. Under these controlled testing environments, developers can experiment with innovative technologies under regulatory oversight, theoretically learning to build safeguards against unintended consequences and bridge the innovation-safety gap.[77]

The AI Act does not establish a uniform EU-wide AI regulatory sandbox nor provide a mechanism by which a broad community of stakeholders can raise sandbox inquiries with national competent authorities. Instead, it mandates national-level sandboxes that may be operated jointly by multiple Member States but remain under national competent authorities' oversight.[78]

Crucially, regulatory sandboxes pose a range of challenges and ethical risks that merit careful consideration. Among the structural challenges are those stemming from the Collingridge dilemma which highlights the difficulty of predicting the consequences of emerging technologies early enough to guide their development. Typically, the establishment of a sandbox coincides with the moment when knowledge about the technology's potential impacts is most limited, which is precisely the point at which regulatory intervention is both most necessary and most uncertain.[79] Approval from a sandbox should not be interpreted as a definitive guarantee of safety, as harmful behaviors or consequences may only emerge in broader deployment contexts or over extended timeframes. This particularly includes issues like 'function creep,' where technologies initially designed for legitimate purposes gradually expand into more intrusive applications.[80] Additional issues include ethics-washing, concerns regarding the independence of supervisory authorities, misplaced trust in organizations or technologies that may not be genuinely trustworthy, and insufficient resources for meaningful oversight and governance of sandbox activities.[81]

Given the temporal and contextual gap between sandbox approval and real-world consequences, post-market monitoring is critically important. Although the

[76] On this distinction see M. Ebers, n 74 above, 686. For more developments, see para IV, no 2 below.

[77] F. Bagni et al, 'White Paper on Regulatory Sandboxes for AI and Cybersecurity' *SSRN paper* (2025), available at https://tinyurl.com/bdvhwtev (last visited 31 January 2026); W.G. Johnson, 'Caught in Quicksand? Compliance and Legitimacy Challenges in Using Regulatory Sandboxes to Manage Emerging Technologies' 17(3) *Regulation and Governance*, 709-725 (2023).

[78] F. Bagni et al, n 77 above, 201.

[79] P. Tõnurist and A. Hanson, *Anticipatory Innovation Governance: Shaping the Future Through Proactive Policy Making* (Paris: OECD, 2020), 25-26, available at https://tinyurl.com/2he7k7rh (last visited 31 January 2026).

[80] B.J. Koops, 'The Concept of Function Creep' 13(1) *Law, Innovation and Technology*, 29-56 (2021), available at https://tinyurl.com/2m7w947z (last visited 31 January 2026).

[81] K. Francis, 'The Need for an Ethical Approach to Regulatory Sandboxes', in F. Bagni et al, n 77 above, 192.

Act allows for this under Art 72(3), its focus on 'known and reasonably foreseeable risks' is particularly problematic. This is especially when considering both the potential gradual expansion of a technology's use beyond its originally intended purpose and the emerging landscape of harmful AI behaviors arising from novel methods of circumventing state-of-the-art ethical safeguards, commonly referred to as 'AI jailbreaks.'[82]

The Fundamental Rights Impact Assessment (FRIA) could have served as a meaningful mechanism to harmonize risk-based and rights-based approaches, aimed at preventing high-risk AI systems from violating fundamental rights.[83] However, its framework is critically undermined by the same limitations found in risk assessments, as it applies selectively, specifically to certain entities (public bodies and entities providing public services) and only to AI systems categorized as high-risk. Once again, the EU AI Act's treatment of high-risk AI systems reveals a regulatory paradigm that implicitly treats AI as a controllable, purely technocratic domain, suggesting a product-centric conceptualization of AI systems. The fact that the FRIA relies more on self-reporting by AI developers and deployers than on robust, independent assessment mechanisms only increases concerns about the adequacy of relying solely on procedural compliance rather than implementing more substantive restrictions.[84]

A critical issue lies also in the nature of fundamental rights and values themselves. How does one quantify the value of privacy, freedom of expression, or human dignity? How does one weigh the potential economic benefits of an AI system against its potential to infringe on fundamental rights? Such decisions are inherently qualitative and involve complex, context-dependent considerations that transcend simple procedural compliance and cost-benefit analyses. As Andy Stirling aptly observes, there is a dramatic need to 'keep it complex,' fostering more plural approaches that can help bridge 'quantitative and qualitative methods,' 'articulate risk assessment and risk management,' and 'reconcile science-based and precautionary appraisal methods.'[85]

This insight resonates strongly. The 'evidence dilemma' stemming from the

[82] M. Russinovich et al, 'Great, Now Write an Article About That: The Crescendo Multi-Turn LLM Jailbreak Attack' (arXiv:2404.01833) (2024), available at https://tinyurl.com/c4ev3fzf (last visited 31 January 2026); A. Rao et al 'Tricking LLMs into Disobedience: Formalizing, Analyzing, and Detecting Jailbreaks', in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (Turin, Italy: ELRA, 2024), available at https://tinyurl.com/5exnfhhc (last visited 31 January 2026).

[83] M. Ebers, n 74 above; G. De Gregorio and P. Dunn, 'The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age' 59(2) *Common Market Law Review*, 473-500 (2022); A. Cosentini et al, 'Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights' *SSRN paper* (2025), available at https://tinyurl.com/2sjeuzfm (last visited 31 January 2026).

[84] A. Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template' 54 *Computer Law & Security Review*, (2024), available at https://tinyurl.com/292nhaws (last visited 31 January 2026).

[85] A. Stirling, 'Keep It Complex' 468 *Nature*, 1029, 1031 (2010). See also M. Kaminski, 'Regulating the Risks of AI' 103 *Boston University Law Review*, 1347 (2023).

rapid pace and unpredictability of AI advancements has not gone unnoticed by the scientific community. The International Scientific Report on the Safety of Advanced AI (29 January 2025) unequivocally warns of risks 'whose likelihood, nature, and timing remain unusually ambiguous.'[86] This stark assessment compellingly underscores the urgent need to reimagine regulation as an uncertainty-focused and contextually sensitive process, rather than a narrow, deterministic pursuit of technological advancement.

As stated in the Report, risk assessment remains an emerging field within the AI safety community, with no fully validated, systematic methods yet available to evaluate the potential severity and likelihood of harm from AI's increasingly autonomous agents. This limitation extends to the possible malicious use of AI in sensitive domains - including smartphone assistants, always-listening voice assistants, healthcare, and legal practice[87] - as well as the risks of potential copyright infringement and violations of data protection laws arising from AI-generated content.[88]

Ultimately, a key takeaway from the report is a profound cautionary note about underestimating the radical asymmetry between human cognitive capabilities and artificial intelligence's dramatically superior processing power. This asymmetry manifests in two critically interrelated dimensions. First, AI's growing capacity to function as autonomous agents capable of independently formulating complex strategies, pursuing multi-step goals, and adapting approaches with computational precision that transcends human cognitive boundaries, and second, the equally concerning potential for humans to manipulate and circumvent even the most sophisticated state-of-the-art ethical safeguards.

The confluence of these factors creates a dual threat landscape where AI's autonomous operational capacity can be strategically exploited or inadvertently misdirected, transforming these systems from controlled technological tools into dynamic, unpredictable entities capable of undermining ethical constraints or human intentions.

Moreover, the persistent 'information gap' between what AI companies know about their systems and what governments and non-industry researchers can access further strengthens this warning.[89] As highlighted in the report, preparing for ambiguous risks requires moving beyond industry-driven frameworks toward a regulatory paradigm that fosters closer collaboration between governments, industry leaders, researchers, and civil society. While this approach does not necessarily bridge the knowledge gap in the lawmaking phase, it would at least ensure that regulatory decisions during implementation are informed by a truly contextually sensitive process, incorporating a wide range of perspectives. We will explore this

[86] Y. Bensho, n 46 above, 100.
[87] ibid 208.
[88] ibid 144.
[89] ibid 22.

issue in greater detail in the following section.

2. The Corporate Epistemic Capture of AI and the Reliance of the EU AI Act on Industry-Driven Standards

It is now well documented that the relationship between major technology companies and the EU AI Act was fraught with tensions and conflicts of interest, with powerful actors deploying considerable resources to steer legislative outcomes, ultimately nudging the balance away from precautionary measures and toward industry-friendly compromises.[90] Yet corporate influence extends far beyond lobbying. Across Europe and globally, Big Tech firms have become the dominant funders of AI research, establishing pervasive financial entanglements with universities, think tanks, and the policy advisory ecosystem itself. This creates what might be termed an epistemic monoculture: a landscape where critical perspectives struggle to gain traction and where the boundaries between independent expertise and industry advocacy grow increasingly porous. The result is not merely regulatory capture in the traditional sense, but something more insidious, that is, a capture of the very knowledge systems upon which democratic governance depends.

With the Act now in force, hopes for more responsible and democratic oversight necessarily shift to the evolving implementation architecture. This implementation phase is critical because the EU AI Act, grounded as it is in the Union's New Legislative Framework (NLF), marries binding legal obligations to a number of bottom-up instruments that complement the Regulation, such as delegated and implementing acts, codes of practice, and harmonized standards.[91] It is indeed within standardization committees and oversight mechanisms that the Act's ultimate character will be forged. Here, abstract legal principles such as fairness, transparency, robustness, and accountability will be translated into operational requirements.

The 'division of labour' implied in this co-regulatory model has proven effective in domains such as product safety, but it poses distinctive challenges in the AI context due to the fact that AI systems pose different risks stemming from their varying levels of autonomy and adaptability, and the domains of their deployment. Once cited in the Official Journal of the European Union, these standards will confer a 'presumption of conformity',[92] meaning that compliance with them is

[90] For a recent overview, see Corporate Europe Observatory, 'Bias Baked In: How Big Tech Sets Its Own AI Standards' (Corporate Europe Observatory Report, 9 January 2025), available at https://tinyurl.com/3db94f9d (last visited 31 January 2026).

[91] D.G. Baeva et al, 'Power to the Standards: Expert Consultation on the Role of Norms and Standards in the European Regulation of Artificial Intelligence' (White Paper, The Center for Trustworthy Artificial Intelligence, 2023), available at https://tinyurl.com/zz8apvrk (last visited 31 January 2026); J. Soler Garrido et al, *Harmonised Standards for the European AI Act* (Seville: European Commission, Joint Research Centre, Report no JRC139430, 2024), available at https://tinyurl.com/3yx4r6us (last visited 31 January 2026).

[92] 'Standard-setting Overview' (16 December 2022), available at https://tinyurl.com/425b8n4s

treated as compliance with the Act's requirements themselves. Understanding this dynamic is thus essential to grasping how and to what extent epistemic capture, established during the legislative phase, threatens to calcify during implementation, potentially hollowing out the Act's protective ambitions from within.

The development of standards on which the future-proof operational effectiveness of the AI Act critically depends is mainly led by two pivotal yet relatively obscure organizations: the European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC). Both present themselves as inclusive, consensus-based organs: private international non-profits operating with 34 national member bodies, claiming the involvement of more than 200,000 technical experts and a wide array of stakeholders-industry, public authorities, academia, consumer and environmental groups, trade unions and SMEs.[93]

Yet behind this official picture of a balanced, transparent, and stakeholder-driven process lies a deeper struggle over power and influence. The very forums tasked with operationalizing the AI Act have become arenas where competing economic, political, and epistemic interests converge, revealing broader concerns about Europe's technological sovereignty and its position in a global race to shape the rules of AI governance. This is well illustrated by several European Commission initiatives designed to limit non-EU entities' influence in standardization processes. A notable example emerged in December 2022 when the European Telecommunications Standards Institute (ETSI), known for its openness to foreign companies, was excluded from the AI Act's draft standardization request.[94]

Can we say that comparable energy was invested in preventing industry capture of the standardization process - so that standards would not simply encode corporate priorities and technological assumptions as if they were objective, value-neutral technical specifications? That seems doubtful.

To be sure, CEN and CENELEC have moved to widen participation, most visibly through JTC 21 on Artificial Intelligence and its Task Group on Inclusiveness-alongside webinars and capacity-building outreach intended to enable civil society engagement. These are important steps as they recognize that the legitimacy of

(last visited 31 January 2026).

93 CEN-CENELEC - European Committee for Standardization and European Committee for Electrotechnical Standardization, available at https://tinyurl.com/4mdfnk8x (last visited 31 January 2026).

94 However, ETSI currently participates in the broader European AI standardization ecosystem through complementary roles. CEN-CENELEC JTC 21 - see below in the text - has been established to lead the bulk of the work, focusing on areas such as risk management, conformity assessment and data governance. ETSI, through its TC SAI group, is developing complementary technical specifications, particularly around robustness, security and testing. The European Commission and its Joint Research Centre (AI Watch) provide coordination and policy oversight, ensuring that technical work remains aligned with the Regulation's objectives. See M. Cantero Gamito, 'The Role of ETSI in the EU's Regulation and Governance of Artificial Intelligence' *Innovation: The European Journal of Social Science Research* (OnlineFirst, 2024), available at https://tinyurl.com/3vymy97j (last visited 31 January 2026).

technical rules depends on broader participation and on translating highly specialized deliberations into terms that non-technical stakeholders can meaningfully understand and contest.[95]

Nevertheless, from a PNS perspective, these procedural gestures mask enduring asymmetries. Industry actors possess far greater resources, technical staff, and institutional footholds to shape committee agendas, draft proposals, and mobilize expert networks. Civil society groups, consumer organizations, trade unions and smaller firms, while formally invited, often lack the time, expertise, and funding needed to participate on equal terms. Transparency gaps, including opaque draft texts, limited access to meeting records, weak conflict-of-interest disclosures, further narrow the space for deliberation about distributional harms, democratic accountability, and long-term societal trade-offs.[96]

These dynamics have come into even sharper focus in recent months with the release of the General-Purpose AI (GPAI) Code of Practice on 10 July 2025 - a significant development in the EU AI Act's efforts to operationalize obligations for GPAI model providers.[97] Although this voluntary framework was crafted through a multi-stakeholder process, it presents indeed the same critical challenges that underscore and perpetuate the broader governance concerns identified in the legislative and standardization phases of the EU AI regime. A broad coalition of civil society organizations indicate that industry actors maintained substantial influence over its drafting and content, with the third draft significantly weakening its approach to systemic risks by shifting responsibility away from model developers and making key risk categories optional.[98]

Against this backdrop, we argue that embedding 'extended peer communities' into AI co-regulation is far more than a bureaucratic tweak; it is an epistemic intervention that brings alternative forms of knowledge to the fore, challenges the dominance of industry-driven narratives, and resists reducing compliance to narrow technical checklists.

[95] https://tinyurl.com/4jxdz5p9.

[96] R. Kilian et al, 'European AI Standards - Technical Standardisation and Implementation Challenges Under the EU AI Act' 16(3) *European Journal of Risk Regulation*, 1038 (2025) and M. Gornet and H. Herman, 'A peek into European standards making for AI: between geopolitical and economic interests' (Hal preprint version), 45 (2024), available at https://tinyurl.com/n7juj5ks (last visited 31 January 2026). These standards have not yet been completed. Initially expected by April or August 2025, the deadline has slipped. Latest projections suggest completion now expected end of 2025 or even into 2026.

[97] Artificial Intelligence (CEN-CENELEC Topics) (1 June 2021) CEN-CENELEC JTC 21, available at https://tinyurl.com/4jxdz5p9 (last visited 31 January 2026).

[98] 'Joint Civil Society Letter Urging the EU Institutions to Protect Fundamental Rights in the Code of Practice for General Purpose AI - Final Draft' (March 2025), available at https://tinyurl.com/4z4hek2r (last visited 31 January 2026).

V. Conclusion: Post-Normal Science as a Framework for Engaging with the Epistemic Infrastructure of EU AI Governance

The EU, with its strong commitment to democratic governance and the precautionary principle, is uniquely positioned to confront the post-normal challenge of AI co-regulation. Doing so, however, requires moving beyond the comfort of technocratic closure and embracing a more demanding, but ultimately more democratic, approach: one that embeds marginalized and alternative forms of knowledge within the co-regulatory framework of the AI Act.

Now that the EU AI Act has entered into force, the decisive arena for responsible and democratic AI governance lies not in the legislative text itself but in the seemingly quiet, procedural machinery of implementation, namely, through standards.

Often presented as neutral tools of rationalization, standards in practice operate as 'technologies of government' in the Foucauldian sense: they shape how actors perceive, categorize, and manage AI-related risks. Their authority stems not only from technical precision but also from their ability to impose a shared language, translating uncertainty into measurable categories and thus producing the very objects of regulation.[99]

Yet this delegation reveals a deep tension. The regulation of an inherently uncertain phenomenon is being transferred to techno-bureaucratic arenas where democratic legitimacy is weak and alternative forms of knowledge - sociological, ethical, civic - are often marginalized or excluded.

The centrality of standardization, it has been observed, faces a paradox:

> 'there is a significant gap between the importance of standardisation and the interest shown in it by academic literature, the scientific world and the general public as a whole.'[100]

Bridging this gap is crucial if Europe's ambition to steer AI's trajectory through a distinctive and responsible model of governance is to be realized.

Critical and independent scholars thus have both an opportunity and a responsibility to engage in building what might be called the epistemic infrastructure of AI governance. What is at stake is not merely the technical detail of conformity assessments or the timing of harmonized standards, but the deeper question of how knowledge about AI systems is produced, validated, and authorized - and which forms of knowing are excluded.

As technologies and practices evolve, so too will these standards. The ongoing process of standard setting will stretch well beyond 2025, increasingly becoming the site where fundamental decisions are made about what can be known about

[99] On the Foucauldian nexus of power/knowledge as constitutive of algorithmic governance, see J. Jarke et al, 'Knowing in Algorithmic Regimes: An Introduction', in Eid eds, *Algorithmic Regimes: Methods, Interactions, and Politics* (Amsterdam: Amsterdam University Press, 2024), 7-34.

[100] M. Gornet and H. Herman, n 96 above, 43.

AI, who has the authority to know it, and what remains beyond regulatory scrutiny. These are not merely technical choices but societal ones, shaping the contours of accountability, justice, and democratic legitimacy in the digital age.

Precisely because standardization is ongoing and revisable, it offers a critical entry point for rethinking governance in post-normal terms. Defining the very frameworks through which AI risks and harms become visible, measurable, and governable constitutes both an epistemic and a political act. Here, critical scholarship has a vital role in advocating for the inclusion of 'uncomfortable knowledge':[101] lived experience data, sociotechnical evidence, and non-Western epistemologies that challenge dominant technical imaginaries of AI.

The question is not simply how to make AI systems 'safe' or 'ethical' according to technical standards, nor merely to ensure a 'human in the loop.' As illustrated in the documentary *Humans in the Loop*, even the most technical dimensions of AI production depend on human judgment, labor, and situated experience.[102] The film exposes how the people whose work makes AI possible are systematically obscured, much like how standardization conceals the normative choices embedded in 'technical' processes. Both reveal a politics of invisibility, whose knowledge and labor are rendered visible, and whose are erased.

PNS calls for 'extended peer communities,' namely, the inclusion of diverse perspectives in decision-making. Yet, the film implicitly challenges this ideal by showing that those already in the loop-data workers, annotators, moderators-rarely hold epistemic or political authority, despite their proximity to AI systems. Their exclusion exemplifies the governance gaps that characterize the current regime of AI co-regulation. The real question, then, is how to ensure that the development and deployment of these technologies remain accountable to democratic processes and serve broader social purposes.

By emphasizing the interdependence of epistemology and governance, PNS calls for adaptive and reflexive regulatory architectures grounded in what has been called epistemic humility,[103] that is, the inclusion of diverse perspectives and experiences in governance processes. Some scholars have gone further, interpreting the 'extended peer community' of PNS as an enactment of standpoint theory, which holds that those most affected by a given risk must play a central role in determining the path forward.[104]

UNESCO's approach to AI governance, exemplified by its RAM and EIA, shows that alternative frameworks are both possible and already being implemented internationally. These initiatives explicitly embrace uncertainty, prioritize engagement with affected communities and civil society, and adopt context-

[101] S. Rayner, n 24 above, 107.

[102] 'Humans in the Loop' (IMDb, 2024), available at https://tinyurl.com/bdds8tem (last visited 31 January 2026).

[103] S. Jasanoff, 'Technologies of humility' n 36 above, 223-244.

[104] D. McQuillan, *Resisting AI: An Anti-Fascist Approach to Artificial Intelligence* (Bristol: Bristol University Press, 2022), 109.

sensitive approaches that ask not only how technologies can be made 'acceptably safe,' but whether they should be deployed at all.

Existing critical studies of AI, while valuable in their own right, remain dispersed across disciplines and subfields, spanning sociology, media studies, law, STS. Yet, they remain only loosely connected within AI governance debates. While it is important to remain mindful of the specificities of the various legitimate voices involved in shaping these frameworks, strengthening the connections among them, particularly around the frameworks through which AI risks and blind spots become visible, is both necessary and promising. Regardless of disciplinary origin, critical AI scholars can find in PNS a shared conceptual and political space for dialogue and a means to challenge epistemic and governance monopolies. This will allow movement beyond both epistemic inertia-the privileging of established ways of knowing-and governance inertia-the resistance to reimagining entrenched institutional arrangements.

In practice, concrete avenues exist to engage with standardization and to contest the epistemic boundaries this process may draw.[105] These arenas are often opaque, resource-intensive, and vulnerable to corporate capture. Yet they remain among the few institutional spaces where the ethical and democratic ambitions of the EU AI Act might still be meaningfully advanced.

Undoubtedly, the issues at stake extend far beyond AI itself. How we choose to govern AI will set important precedents for how democratic societies confront future technological disruptions in an era of accelerating change. Ultimately, this process will prompt a broader reflection on how ostensibly technical arenas shape the material and epistemic foundations of rights and freedoms in the age of AI and on how knowledge itself becomes both a source and an outcome of law-making.

[105] For instance, scholars can engage substantively in national mirror committees to shape CEN and CENELEC drafts within member state processes; respond to public consultations to submit evidence-based critiques; publish alternative technical papers - on metrics of fairness, bias detection methodologies, or robustness datasets - that provide reference points for future standards; and build coalitions with civil society organizations, consumer groups, and trade unions to amplify calls for inclusivity and pluralism in standard-setting.