LAW AND REGULATION

# Calibrating Algorithmic Accountability: Risk-Based Regulation, Enforcement Capacity and Deferral in European Artificial Intelligence Governance

**Murager Abylassimov**

Karagandy National Research University named after Academician E. A. Buketov, Karaganda, Kazakhstan
ORCID 0000-0002-7730-7719 · muragerabylasimov@gmail.com

PUBLISHED 31 MARCH 2025

CITE THIS ARTICLE Abylassimov, M. (2025). Calibrating Algorithmic Accountability: Risk-Based Regulation, Enforcement Capacity and Deferral in European Artificial Intelligence Governance. Journal of Digital Economy, Law and Policy, 1(1), 5–13.

ABSTRACT

Risk-based regulation has become the dominant paradigm for governing artificial intelligence, and Regulation (EU) 2024/1689 (the AI Act) is its most complete legislative expression. This article argues that the paradigm's principal weakness lies not in its substantive design but in the mismatch between the obligations it imposes and the institutional capacity available to enforce them. Drawing on doctrinal and comparative-institutional analysis of the AI Act's staged application, the unresolved position of the proposed AI Liability Directive, and parallel developments in the United States, the United Kingdom, China and the Council of Europe, the article identifies three structural pathologies. The first is a deferral dynamic: because risk-based regimes concentrate cost in ex ante conformity assessment, and that cost is borne before any harm materialises, compliance deadlines become politically renegotiable – a pressure already visible in the distance between the August 2026 date for Annex III high-risk obligations and the state of the harmonised standards on which conformity assessment depends. The second is a remedial asymmetry: with the AI Liability Directive stalled, public supervision has no matching private enforcement channel, so deterrence rests almost entirely on administrative capacity. The third is a capacity gradient: the transaction costs of risk-based regulation fall disproportionately on jurisdictions with thin supervisory institutions, which makes wholesale transplantation of the European model inadvisable for many emerging economies. The article proposes a capacity-indexed sequencing model in which transparency duties, procedural rights and sectoral supervision precede general-purpose conformity assessment.

KEYWORDS artificial intelligence regulation; algorithmic accountability; EU AI Act; risk-based regulation; regulatory capacity; digital governance; liability

JEL K23; K24; L51; O38

IRSTI 10.19.00

UDC 34:004.8

## 1. Introduction

Within a decade, artificial intelligence has moved

from a specialised research field to a general-purpose technology embedded in credit scoring, recruitment, medical triage, border management, criminal justice and public administration. The regulatory response has been correspondingly rapid. Between 2021 and 2024, dozens of jurisdictions adopted binding or quasi-binding instruments addressing automated decision-making, and the European Union completed the first comprehensive horizontal statute in the field. Regulation (EU) 2024/1689 – the AI Act – entered into force on 1 August 2024 and applies in stages from February 2025.

The intellectual foundation of this response is risk-based regulation: the proposition that regulatory intensity should be proportionate to the expected harm of a given application, so that scarce enforcement resources are directed to the uses that matter most (Black, 2010; Baldwin & Black, 2016). The AI Act operationalises this by sorting systems into prohibited practices, high-risk systems, limited-risk systems subject to transparency duties, and minimal-risk systems, and by adding a parallel tier for general-purpose AI (GPAI) models. The architecture is elegant, and it has proved influential: it is visible in the drafting of Brazil's PL 2338/2023, in the Council of Europe's Framework Convention on Artificial Intelligence, and in the vocabulary of national AI strategies across Asia, Africa and the post-Soviet space.

Yet the paradigm is under strain before its central obligations have taken effect. The proposed AI Liability Directive – the private-law limb of the original 2022 package – has made no substantive progress in the Council, and its withdrawal is openly canvassed. The harmonised standards on which conformity assessment depends are behind the schedule set for them, and national market surveillance authorities are being designated with resources that bear little relation to the mandate they are given. Outside the Union the picture is no steadier: Colorado's Artificial Intelligence Act, the closest state-level analogue to the European model, was enacted in May 2024 with an effective date of 1 February 2026 and is already the subject of amendment proposals; the United Kingdom has declined to legislate horizontally; and federal policy in the United States has oscillated between administrations without producing a statute. The direction of travel is unmistakable: the ambition expressed in statutory text is being trimmed at the point of application.

This article asks why. The conventional explanation is political: industry lobbying, competitiveness anxiety, and a change in the ideological climate. That explanation is not wrong, but it is incomplete, because it does not tell us why risk-based AI regulation should be more susceptible to such pressure than, say, data protection or financial regulation, which survived comparable pressure. The argument advanced here is institutional. Risk-based AI regulation has three design features that make its obligations unusually renegotiable: it front-loads cost before harm is observable; it depends on technical standards and conformity infrastructure that legislatures do not control; and – in the European case – it was stripped of the private enforcement channel that would otherwise have made deferral of public obligations less consequential.

The article proceeds as follows. Section 2 develops a conceptual framework distinguishing four components of algorithmic accountability. Section 3 reviews the relevant literature. Section 4 sets out the methodology. Section 5 presents the analysis in five parts: the architecture of the AI Act, the deferral dynamic, the remedial asymmetry created by the withdrawal of the AI Liability Directive, enforcement capacity as a binding constraint, and a comparative survey of alternative regulatory trajectories. Section 6 discusses the findings and derives three propositions. Section 7 develops policy implications, with particular attention to jurisdictions outside the OECD core. Section 8 notes limitations, and Section 9 concludes.

## 2. Conceptual Framework: What Algorithmic Accountability Requires

"Accountability" is used loosely in the AI policy literature, often as a synonym for transparency or for the availability of an explanation. Following Bovens (2007), accountability is better understood as a relation: an actor is obliged to explain and justify conduct to a forum, the forum can pose questions and pass judgement, and the actor may face consequences. Applied to algorithmic systems, this yields four components, each of which must be independently satisfied for the relation to exist.

Legibility. The forum must be able to know that an automated system was used, what it was used for, and on what basis it produced a given output. This is the domain of documentation duties, model cards, technical files, logging obligations and the contested "right to an explanation" (Wachter et al., 2017; Selbst & Powles, 2017). Legibility is a

precondition, not accountability itself; a perfectly documented system to which no consequence attaches is not an accountable one.

Attribution. The forum must be able to identify who is answerable. This is unusually difficult in AI value chains, where a foundation model developer, a fine-tuner, a system integrator, a deployer and a distributor may each contribute to the outcome. The AI Act's provider/deployer distinction and its GPAI value-chain provisions are attempts to solve this, but the allocation remains contested where a downstream actor substantially modifies an upstream model.

Contestability. The affected person must have a procedural route to challenge an outcome. This is the function of Article 22 GDPR, of the AI Act's right to explanation of individual decision-making (Article 86), and of sectoral rights in credit, employment and social security law. Contestability is the component most often assumed and least often instrumented (Almada, 2019; Kaminski & Urban, 2021).

Consequence. There must be a realistic prospect that failure produces a cost – administrative penalty, civil liability, market exclusion, or reputational sanction. Without this component the first three collapse into a documentation exercise.

The framework matters because risk-based regulation, as implemented in the AI Act, is disproportionately strong on legibility and attribution, moderately strong on contestability, and structurally weak on consequence. The regime's penalty ceilings are high – up to €35 million or 7% of worldwide annual turnover for prohibited practices, and €15 million or 3% for most other infringements – but ceilings are not deterrence; realised enforcement is. The empirical question is therefore whether the institutional apparatus exists to convert paper obligations into realised consequence. Section 5 argues that it does not yet, and that the pressure already visible on the 2026 compliance dates is the symptom of that gap.

## 3. Literature Review

Three strands of scholarship converge on this problem.

The first is the regulatory-governance literature on risk-based and meta-regulation. Black (2010) and Baldwin and Black (2016) show that risk-based frameworks depend on the regulator's ability to score risk reliably and to reallocate resources accordingly; where the scoring is contested or the resources fixed, risk-based regulation degenerates into formal categorisation without differential supervision. Coglianese and Lehr (2017) extend the point to machine learning specifically, noting that the epistemic opacity of learned models undermines the regulator's own capacity to verify claims made in conformity documentation. Gunningham and Sinclair's work on smart regulation supplies the complementary insight that regulatory instruments function as portfolios: weakening one instrument alters the effectiveness of the others.

The second strand is the algorithmic accountability literature in law and science-and-technology studies. Citron and Pasquale (2014) established the procedural-due-process framing for automated scoring; Pasquale (2015) supplied the "black box" diagnosis; Barocas and Selbst (2016) demonstrated that disparate impact in machine learning arises from ordinary modelling choices rather than from malice, which has the important regulatory implication that intent-based liability rules will systematically under-capture algorithmic harm. Kaminski (2019) analyses the GDPR as a hybrid of individual rights and systemic (collaborative) governance, and argues that the two limbs are complements: individual transparency rights supply the information that makes systemic oversight tractable, while systemic oversight supplies the enforcement capacity that individual rights lack. Veale and Zuiderveen Borgesius (2021) offer the most influential early critique of the AI Act proposal, characterising it as a product-safety statute wearing the vocabulary of fundamental rights, and predicting that reliance on harmonised standards would shift substantive policymaking into standardisation bodies with limited democratic accountability – a prediction that the delays in CEN-CENELEC standardisation work have largely borne out.

The third strand concerns liability and remedies. Wagner (2018) and Ebers (2021) analyse the fit between product-liability doctrine and software; the recurring finding is that the defect concept, the development-risk defence and the burden of proof each operate against claimants in algorithmic cases. Hacker (2023) argues that the AI Act's public-law architecture presupposes a functioning private-law backstop. The stalling of the AI Liability Directive makes this argument testable in a way it was not when written.

What the literature does not yet supply is an

account of why compliance deadlines in risk-based AI regimes are systematically unstable, and of what this instability implies for jurisdictions considering transplantation of the model. That is the contribution attempted here.

## 4. Methodology and Data

The study uses doctrinal legal analysis combined with comparative institutional analysis. It is qualitative and interpretive; it does not test a statistical hypothesis, and no original quantitative dataset is generated.

The doctrinal component reconstructs the operative provisions of Regulation (EU) 2024/1689, together with Directive (EU) 2024/2853 on liability for defective products, Article 22 and Chapter III GDPR, and the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. Sources are limited to authoritative texts: the Official Journal, Commission communications and work programmes, published decisions, and the official documentation of the national authorities discussed.

The comparative component follows a most-similar-systems logic across four jurisdictions selected for variation on the dependent variable (regulatory intensity) while holding constant high AI-development capacity: the European Union (comprehensive horizontal statute), the United Kingdom (sectoral, principles-based, regulator-led), the United States (fragmented, state-led, with active federal preemption pressure), and China (targeted, service-specific, ex ante licensing). Four institutional dimensions are coded for each: the locus of obligation-setting, the mechanism of conformity verification, the availability of private remedies, and the observed stability of compliance deadlines.

The analysis covers developments to 31 January 2025. Two limitations follow from the design. First, because much of the regime has not yet reached its application date, claims about enforcement outcomes are necessarily prospective rather than observed. Second, the assessment of institutional readiness rests on published designations, budget allocations and standardisation work programmes, which are incomplete and are reported at different intervals across Member States.

## 5. Analysis

### 5.1 The Architecture of Risk-Based AI Regulation

The AI Act's architecture rests on three moves. It classifies by use rather than by technology, so that the same model attracts different obligations depending on deployment context. It borrows the New Legislative Framework from EU product-safety law, so that compliance is demonstrated through conformity assessment, harmonised standards, CE marking and market surveillance rather than through case-by-case adjudication. And it superimposes a model-layer regime on GPAI, recognising that use-based classification alone cannot reach systems whose downstream applications are indeterminate at the point of development.

Each move carries a cost. Use-based classification requires the regulated entity to characterise its own deployment context, which invites strategic classification at the margin – a system described as "decision support" rather than "decision-making" may fall outside Annex III on a contestable reading. Reliance on the New Legislative Framework transfers substantive specification to standardisation bodies; where the standards are late, the primary obligation becomes practically unverifiable, and the regulator faces a choice between enforcing against an undefined benchmark and postponing. The GPAI layer creates an allocation problem between model providers and downstream deployers that the statute addresses through information-flow duties rather than through a clear liability rule.

The staged application timetable was designed to manage these costs. Prohibitions and AI-literacy duties applied from 2 February 2025; GPAI obligations from 2 August 2025; high-risk obligations follow on 2 August 2026 for Annex III stand-alone systems and 2 August 2027 for Annex I systems embedded in regulated products. Governance and penalty provisions were phased alongside. On paper, the sequencing is rational: the cheapest and least contested obligations first, the most infrastructure-dependent last.

### 5.2 The Deferral Dynamic: Why Ex Ante Deadlines Slip

That sequencing is rational only if the infrastructure it presupposes arrives on time, and there is little sign that it will. Conformity

assessment under Annex III depends on harmonised standards that the European standardisation bodies have not delivered, on notified bodies that have not been accredited in the numbers required, and on national market surveillance authorities that in most Member States have been neither designated nor funded. The August 2026 date was fixed on an assumption about readiness that the standardisation record has not borne out. The dynamic that follows is general rather than European. In an ex ante conformity regime, compliance cost is incurred by every regulated entity before any harm is observed, while the benefit is diffuse and counterfactual. That distribution creates a standing, well-organised constituency for postponement and no comparable constituency for adherence. Where the regulator also knows that the assessment infrastructure is incomplete, enforcing the original date would expose it to the charge of penalising firms for the absence of standards it was itself responsible for procuring. The rational move for both sides is delay, and the instrument for delay – a targeted amendment to the application dates – is legislatively cheap.

Not every layer is equally exposed. Obligations that a central authority can enforce directly against a small number of identifiable actors – the Article 5 prohibitions, the Article 50 transparency and labelling duties, and supervision of general-purpose model providers – depend on neither notified bodies nor harmonised standards, and there is no comparable pressure to move them. The exposure is concentrated precisely where conformity infrastructure is load-bearing.

The pattern this predicts is analytically significant. What is exposed is precisely the layer that depends on conformity infrastructure – harmonised standards, notified bodies, technical documentation review. What is insulated is the layer that a central authority can enforce directly against a small number of identifiable actors: prohibitions, labelling duties, and supervision of frontier model providers. A regime that trims in this pattern is not retreating from regulation but reallocating toward what its institutional apparatus can actually deliver. Read this way, a postponement of the high-risk layer would be evidence for the capacity hypothesis rather than merely for the lobbying hypothesis.

The deferral dynamic has a general form. In an ex ante conformity regime, compliance cost is incurred by all regulated entities before any harm is observed, and the benefit – avoided harm – is diffuse, probabilistic and invisible when the regime works. The political economy is therefore asymmetric: the constituency for postponement is concentrated, informed and organised; the constituency for timely application is diffuse and cannot point to a specific injury caused by delay. Ex post liability regimes have the opposite profile, which is why deadlines in tort-based regimes are rarely renegotiated. The implication is that any risk-based regime lacking a liability backstop should be expected to experience deadline slippage as a structural feature, not as an accident.

### 5.3 The Remedial Asymmetry: Life After the AI Liability Directive

The original 2022 package paired the AI Act with a proposed AI Liability Directive (AILD), which would have introduced a rebuttable presumption of causation and a disclosure mechanism for evidence held by defendants in fault-based claims involving AI. The proposal has made no substantive progress in the Council since 2022, and the absence of foreseeable agreement has made its withdrawal a live possibility.

The consequence, if the proposal does not survive, is that private redress for AI-related harm in the Union runs through three imperfect channels. The revised Product Liability Directive (EU) 2024/2853, which Member States must transpose by 9 December 2026, is the most significant: it brings software, including AI systems and their updates, expressly within the product concept; recognises loss of data as recoverable damage; introduces disclosure of evidence and rebuttable presumptions of defectiveness and causation where technical or scientific complexity makes proof excessively difficult; and narrows the development-risk defence. But it is a strict-liability regime for defective products, and it covers death, personal injury, property damage and data loss – not the discriminatory, dignitary, reputational and economic harms that dominate algorithmic-harm litigation. A rejected loan applicant, an unfairly screened job candidate or a wrongly flagged welfare recipient falls largely outside it.

The second channel is national fault-based liability, where the claimant must establish breach, causation and damage against an opaque technical system without the AILD's evidentiary presumptions. The third is data-protection law: Article 82 GDPR provides compensation for material and non-material damage, and Article 22 supplies a

substantive constraint on solely automated decisions producing legal or similarly significant effects. Data-protection litigation has therefore become the principal private-law vehicle for algorithmic grievance in Europe, which is doctrinally awkward – it channels claims about discrimination and procedural unfairness through a statute designed to govern the processing of personal data.

The asymmetry matters for the argument. Where a strong private enforcement channel exists, postponing public obligations is costly to regulated entities: the underlying duty of care survives, and delay in public compliance does not immunise against claims. Where it does not, postponement is close to costless. The withdrawal of the AILD and the deferral of high-risk obligations are thus not independent events; the first materially lowered the price of the second.

### 5.4 Enforcement Capacity as the Binding Constraint

The AI Act's enforcement architecture is polycentric. Market surveillance authorities designated by Member States supervise high-risk systems; notifying authorities oversee conformity assessment bodies; the AI Office supervises GPAI at Union level; the European Artificial Intelligence Board coordinates; and sectoral regulators retain jurisdiction where AI is embedded in already-regulated products and services.

Three capacity problems follow. The first is scarcity of technical expertise. Verifying a conformity claim about a high-risk system requires personnel who can read model documentation, interrogate evaluation protocols and assess data-governance practice. This skill set is scarce and expensive, and public authorities compete for it with the entities they supervise. The second is institutional fragmentation: where twenty-seven Member States designate authorities of widely differing capability, the effective standard becomes the one applied in the jurisdiction where the provider establishes itself, replicating the forum-shopping pathology that has hampered the GDPR's one-stop-shop mechanism. The third is standardisation dependency: until harmonised standards are published, providers cannot rely on the presumption of conformity and authorities cannot benchmark compliance, which converts an obligation to comply into an obligation to negotiate.

Centralising supervision of general-purpose models in the AI Office is a sensible partial answer, because frontier model providers are few, identifiable and well-resourced, which makes centralised supervision feasible. But it does not address the long tail: the thousands of deployers of high-risk systems in employment, education, credit and public administration, where the harms accrue and where supervisory capacity is thinnest.

### 5.5 Comparative Regulatory Trajectories

The United Kingdom declined a horizontal statute in favour of a principles-based, regulator-led model in which existing sectoral regulators apply five cross-cutting principles within their remits. The approach conserves scarce expertise by locating it where domain knowledge already exists, and avoids the standardisation bottleneck. Its weakness is coverage: harms occurring outside any regulator's remit have no home, and the principles lack direct legal effect.

The United States presents the sharpest contrast. There is no federal statute, executive policy has oscillated between administrations, and the substantive action is at state level. Colorado's Artificial Intelligence Act, enacted in May 2024, is the state statute closest to the European model: it imposes impact assessments and duties of reasonable care to avoid algorithmic discrimination on developers and deployers of high-risk systems, and takes effect on 1 February 2026. It was signed with an accompanying statement urging amendment before that date, and a task force was convened to propose changes. The obligations most likely to be revisited are the assessment duties rather than the notice duties – which is the pattern this article predicts. Where an ex ante assessment regime meets a legislature without the capacity to support it, the assessment layer is what gives way, and what survives is disclosure: cheaper to comply with and cheaper to enforce.

China regulates neither horizontally nor by principle but by service type, imposing filing, security-assessment and content-labelling duties on defined categories of provider – algorithmic recommendation services, deep synthesis services, and public-facing generative AI services – with labelling requirements for synthetically generated content at draft stage. The model achieves rapid, enforceable implementation because obligations attach to identifiable licensed intermediaries and because enforcement capacity is concentrated. It achieves this at the cost of individual contestability: the accountability relation runs from provider to state, not from provider to affected person.

The Council of Europe Framework Convention, opened for signature on 5 September 2024, occupies a different register: it binds parties to ensure that activities within the AI lifecycle are consistent with human rights, democracy and the rule of law, and requires the availability of effective remedies, but leaves the choice of instruments to the parties. For jurisdictions lacking the capacity to operate a conformity-assessment infrastructure, it offers a way to commit to the outcomes of algorithmic accountability without adopting the EU's particular means.

Coded on the four institutional dimensions, the comparison yields a consistent pattern. Regimes that locate obligation-setting in standardisation bodies and verification in ex ante conformity assessment (EU) exhibit deadline instability. Regimes that locate obligation-setting in existing sectoral regulators (UK) or in licensing conditions attached to identifiable intermediaries (China) exhibit deadline stability but narrower coverage or weaker individual remedies respectively. Regimes with contested constitutional foundations and no settled locus of obligation-setting (US states) exhibit both instability and coverage loss.

## 6. Discussion

Three propositions follow.

Proposition 1: Deadline stability in AI regulation is a function of enforcement-channel diversity, not of statutory ambition. The EU's high-risk regime is the most ambitious in the world and has the least stable timetable. Its Article 50 transparency duties, which require no conformity infrastructure and can be verified by inspection of outputs, applied on schedule. The variable that distinguishes them is not importance but verifiability. Legislatures designing AI regimes should therefore treat the availability of a verification method as a design constraint, not as an implementation detail to be resolved later by standardisation.

Proposition 2: Public and private enforcement are complements, and removing one raises the political price of maintaining the other. The withdrawal of the AILD did not merely leave a gap in private redress; it removed the shadow of liability that would have made deferral of public obligations less attractive. Where a claimant can sue regardless of the regulator's timetable, postponing administrative deadlines buys the regulated entity less. This suggests that the sequencing debate in emerging jurisdictions – "should we legislate liability or supervision first?" – is misconceived; a supervision-only regime is unstable in a way that a liability-plus-supervision regime is not.

Proposition 3: The transaction costs of risk-based regulation scale with the number of regulated entities, not with the level of risk. Conformity assessment imposes a broadly fixed cost per system, which means the aggregate supervisory burden depends on how many systems fall within scope. Regimes that draw the high-risk perimeter widely therefore create burdens that grow independently of the harm they avert. This is the structural reason why the long tail of deployers – rather than the small set of frontier developers – is where enforcement will first fail, and it argues for concentrating ex ante duties on a narrow set of applications while relying on ex post mechanisms for the remainder.

These propositions do not amount to a case against risk-based regulation. The classification logic is sound, and the alternative – technology-neutral general law applied case by case – has demonstrably failed to produce timely responses to algorithmic harm. The argument is rather that risk-based regulation is a capacity-intensive instrument that has been adopted in several jurisdictions on the assumption that capacity would follow the statute. The record so far – standards behind schedule, supervisory authorities undesignated or unfunded, and a private-law limb that has not advanced – suggests it does not.

## 7. Policy Implications

For the European Union, the implications are narrow but consequential. The window before August 2026 is short, and it will be wasted unless it is used to complete harmonised standards, fund national market surveillance authorities at levels commensurate with their mandate, and address the remedial gap the AI Liability Directive was meant to fill – whether through a targeted instrument on evidentiary presumptions, through amendment of the Product Liability Directive's damage heads, or through explicit procedural rights of contestation attached to Article 86. Without that, postponement of the high-risk layer is a realistic prospect, and each successive postponement would erode the credibility on which the regime's extraterritorial influence depends.

For jurisdictions outside the OECD core – including the Central Asian, South Caucasus and

South-East European economies now drafting AI legislation – the implications are more far-reaching. The temptation to transpose the AI Act's structure is strong: it is available, it is comprehensive, and adopting it signals regulatory alignment to investors and to the European market. But transposing a conformity-assessment regime without notified bodies, accredited testing laboratories, harmonised standards or specialised supervisory staff produces a statute that cannot be enforced, and an unenforceable statute is worse than a narrower one because it teaches regulated entities that compliance is optional.

A capacity-indexed sequencing model would instead proceed in three stages. Stage one consists of instruments that require no technical infrastructure: mandatory disclosure that an automated system is in use, public registration of AI systems used by public authorities in consequential decisions, labelling of synthetic media, and prohibitions on a narrow set of practices that can be verified by observation rather than by audit. Stage two adds procedural rights – a right to human review, a right to a statement of reasons, and a shifted burden of proof in discrimination claims where an automated system contributed to the decision – which locate enforcement in courts and tribunals that already exist rather than in agencies that do not. Stage three, undertaken only where a testing and certification infrastructure has been built or where mutual recognition with a partner jurisdiction is available, introduces ex ante conformity assessment for a deliberately narrow set of high-risk applications, prioritising public-sector uses where the state is simultaneously deployer and supervisor and can therefore act without building external capacity.

This sequencing has an additional advantage: each stage generates the information required to design the next. Registration obligations reveal where AI is actually deployed; procedural-rights litigation reveals which harms are prevalent and which evidentiary obstacles bind; only then is a legislature in a position to draw a high-risk perimeter on evidence rather than on transplantation.

## 8. Limitations and Future Research

Three limitations qualify the analysis. First, the study is doctrinal and comparative-institutional; it does not measure enforcement outcomes, because the principal obligations analysed have not yet reached their application dates. Second, the assessment of institutional readiness rests on published designations, budgets and standardisation work programmes, which are incomplete and unevenly reported across Member States. Third, the four-jurisdiction comparison, while chosen for variation on the dependent variable, cannot exclude confounding by legal family, market size or the domestic presence of major AI developers.

Three research directions follow. The first is empirical measurement of deadline stability across a larger sample of digital-regulation instruments, which would allow the deferral dynamic proposed here to be tested rather than illustrated. The second is quantitative estimation of the compliance cost of conformity assessment per high-risk system, disaggregated by firm size, which would permit the capacity gradient to be specified rather than asserted. The third is comparative study of algorithmic-harm litigation under data-protection, product-liability and anti-discrimination law, to establish which doctrinal route in fact delivers remedies to claimants.

## 9. Conclusion

The AI Act remains the most consequential instrument of algorithmic governance yet enacted, and its influence on legislative drafting far beyond Europe is not in doubt. But the position as its central obligations approach – a private-law limb that has not advanced, harmonised standards behind schedule, market surveillance authorities undesignated or unfunded, and the closest United States analogue already under amendment pressure before it takes effect – suggests that the regime's binding constraint is institutional rather than legislative.

The argument of this article is that these events share a common structure. Risk-based regulation front-loads cost, defers benefit, and depends on verification infrastructure that legislatures do not directly control; where it is not paired with a private enforcement channel, its deadlines are politically renegotiable by design. Accountability, on the framework developed in Section 2, requires legibility, attribution, contestability and consequence. European AI regulation has built the first two well, the third partially, and the fourth barely. Closing that gap requires less new statutory text and more of what is unglamorous: funded supervisors, completed standards, and a private-law

route by which a person harmed by an automated decision can obtain a remedy without waiting for a regulator's timetable.

For jurisdictions still drafting, the lesson is not that the European model should be rejected but that it should be sequenced. A short statute whose obligations can be enforced tomorrow does more for algorithmic accountability than a comprehensive one whose central provisions arrive, if at all, years after enactment.

## REFERENCES


Almada, M. (2019). Human intervention in automated decision-making: Toward the construction of contestable systems. In Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law (pp. 2-11). Association for Computing Machinery.

Baldwin, R., & Black, J. (2016). Driving priorities in risk-based regulation: What's the problem? Journal of Law and Society, 43(4), 565-595.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671-732.

Black, J. (2010). Risk-based regulation: Choices, practices and lessons being learnt. In OECD, Risk and regulatory policy: Improving the governance of risk (pp. 185-224). OECD Publishing.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447-468.

Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. Washington Law Review, 89(1), 1-33.

Coglianese, C., & Lehr, D. (2017). Regulating by robot: Administrative decision making in the machine-learning era. Georgetown Law Journal, 105(5), 1147-1223.

Ebers, M. (2021). Liability for artificial intelligence and EU consumer law. Journal of Intellectual Property, Information Technology and E-Commerce Law, 12(2), 204-220.

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

Kaminski, M. E. (2019). Binary governance: Lessons from the GDPR's approach to algorithmic accountability. Southern California Law Review, 92(6), 1529-1616.

Kaminski, M. E., & Urban, J. M. (2021). The right to contest AI. Columbia Law Review, 121(7), 1957-2048.

Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.

Selbst, A. D., & Powles, J. (2017). Meaningful information and the right to explanation. International Data Privacy Law, 7(4), 233-242.

Smuha, N. A. (2021). From a “race to AI” to a “race to AI regulation”: Regulatory competition for artificial intelligence. Law, Innovation and Technology, 13(1), 57-84.

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the draft EU Artificial Intelligence Act. Computer Law Review International, 22(4), 97-112.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), 76-99.

Wagner, G. (2018). Robot liability. In S. Lohsse, R. Schulze, & D. Staudenmayer (Eds.), Liability for artificial intelligence and the internet of things (pp. 27-62). Nomos.

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. Regulation & Governance, 12(4), 505-523.


## LEGAL INSTRUMENTS AND CASES


Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation), OJ L 119, 4.5.2016, pp. 1-88.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L, 12.7.2024.

Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products, OJ L, 18.11.2024.

Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225, Vilnius, 5 September 2024.