WOLFGANG SCHULZ

Co-Regulation of Frontier AI:

Conditions of Functioning and Conditions of Reliance

HIIG DISCUSSION PAPER SERIES 2026-09

ABSTRACT

Stakeholders in the United States are weighing an industry-funded body, under federal supervision, that would verify before release that developers have tested their models for a small set of catastrophic capabilities and applied their own published safeguards. What is to be verified is therefore compliance with frameworks the developers have written themselves, the benchmarks for capabilities being left to the new structure to set later. The proposals are taken here as a starting point. The question is whether the assessment of frontier models can be co-regulated at all: what the research on co-regulation in media and communications says about the conditions under which such arrangements work, what follows from those conditions for the place of independent research, and what makes the findings of such a body usable in other regulatory orders.

KEYWORDS

Co-regulation, Frontier AI models, Independent research, Auditing, Regulatory transfer

CITATION

Schulz, W.. (2026). Co-Regulation of Frontier AI: Conditions of Functioning and Conditions of Reliance. HIIG Discussion Paper Series 2026-9. 18 pages. https://doi.org/10.5281/zenodo.22772120.

AUTHOR INFO

Prof. Dr. Wolfgang Schulz is Research Director at the HIIG. He dedicates himself intensively to research on Communications Constitutional Law, Media Law, and Internet Governance, with a particular focus on intermediaries, algorithmic decision-making and the ethics of digitisation. To be fruitful in solving complex problems, he relies on the functional comparison of different national approaches. An important role in this regard is played by the transfer of his research results to politics at all levels (regional, national, EU, Council of Europe, UNESCO). He studied law and journalism in Hamburg and holds a professorship in Media Law and Public Law at the University of Hamburg. He is also Director of the Leibniz Institute for Media Research | Hans-Bredow-Institut (HBI) and a Principal Investigator in the DFG Research Unit (FOR 5656) "Communicative AI: The Automation of Societal Communication" (ComAI).

The author would like to thank Dr Martin Fertmann, Jonathan Nörz, Moritz Wiechert and Valentin Peddinghaus for their invaluable advice and support in preparing this manuscript.

CONTENTS

1 INTRODUCTION

The case for regulating frontier models does not rest on apocalyptic assumptions.[1] Some of the risks are already observable. Capabilities in the cyber domain that were recently thought distant have been reached, and providers have themselves disclosed incidents in which agents left the environments set for them. Other effects are less spectacular and reach further, among them shifts in labour markets and changes in how people receive information.

Since those effects cross borders while the systems producing them are built in a handful of states, the question is international from the outset. Proposals for a global arrangement are accordingly on the table. The German federal minister for digital transformation and government modernisation argued in September 2026 for an international oversight body for artificial intelligence on the model of the International Atomic Energy Agency, and the United Nations General Assembly established an independent scientific panel and a Global Dialogue on AI Governance in 2025, the panel presenting its preliminary report on 1 July 2026 and the Dialogue meeting in Geneva in the same month.[2] How far the model of that agency carries is a separate question.[3] Two things matter for what follows: no arrangement of this kind can be built without the industry, because the capabilities of a model are known first to the laboratory that trained it, and connecting to the structures now taking shape in the United States is more promising than opening a wholly new path.

The formula of controlling AI models covers risks with little structurally in common: disruption of labour markets, effects on democratic opinion formation, attacks on critical infrastructure and, at the extreme, assistance in building biological weapons. Some of these risks are inherent in the systems; others arise from misuse by bad actors. They differ in addressee, in time horizon and above all in testability. Labour-market effects cannot be established at the model, and effects on opinion formation depend on national constitutional traditions and belong to media and platform law rather than to a testing regime. Only the narrow band of security-relevant capabilities resembles non-proliferation in structure. What follows concerns primarily that narrow band of security-relevant capabilities alone, where the American proposals also sit. The consequence for institutional design is taken up in section 5: no single process fits all relevant risks, but they are better coordinated on one regulatory platform than distributed across separate ones.

The American debate has reached the same institutional question from the other end. Instead of a body "above" states it proposes one between the federal government and the firms, and it is further advanced, because the addressees have made the proposals themselves. Published in June 2026, they converge on the object, a short list of catastrophic capabilities, and on a layered structure of developer self-assessment, external check and federal supervisor, but not on the institutional form. The most detailed proposes a Frontier AI Regulatory Organization: industry-funded, supervised by a federal agency, governed by industry and independent members, setting safety and security standards, defining capability benchmarks in

[1] Such fears are now voiced ever more forcefully from within the industry itself. https://www.theatlantic.com/ideas/2026/09/ai-destroy-world-apocalypse/688575/

[2] Proposal of the German Federal Minister for Digital Transformation and Government Modernisation, K. Wildberger, in an interview with POLITICO of 3 September 2026, "Wildberger fordert globale KI-Behörde nach Vorbild der Atomaufsicht" https://www.politico.eu/article/karsten-wildberger-fordert-globale-ki-behorde-nach-vorbild-der-atomaufsicht/ accessed 16 September 2026; UNGA Res. A/RES/79/325 of 26 August 2025, establishing the Independent International Scientific Panel on AI and the Global Dialogue on AI Governance; Independent International Scientific Panel on AI, Preliminary Report, 1 July 2026; first session of the Dialogue, Geneva, 6 and 7 July 2026, see United Nations, "Global Dialogue on AI Governance" https://www.un.org/global-dialogue-ai-governance/en ; accessed 16 September 2026. The differentiation in the next paragraph is the author's.

[3] Its safeguards attach to a physical bottleneck, since fissile material can be accounted for and installations entered, whereas models are copyable and the knowledge to produce them diffuses; and it rests on a shared written objective, which has no counterpart here, where the same technology counts at once as an economic and a security advantage. What transfers is infrastructure rather than control: a technical secretariat with its own expertise, standardised reporting channels and common test procedures.

the cyber and CBRN domains and overseeing annual audits before models reach the market, with the Financial Industry Regulatory Authority and the electric reliability organisation named as templates. A second is closer to enforced self-regulation, and a third proposes a designated federal agency able to require evaluations and, in severe cases, to block deployment, with mandatory third-party testing, an arrangement not far from the structure of the Digital Services Act. None of them says much about who, apart from the developers and the supervisor, takes part: those affected by a deployment, the civil society organisations that would speak for them and the research community that would have to supply the measures appear in the documents, but not as part of the procedure, which is the gap that platform governance spent the years before 2013 identifying.[4] The most detailed of them at least envisages that the findings produced will eventually be honoured in other jurisdictions.

The framing of a middle way between over-regulation and no regulation misdescribes the alternatives, because there is no unregulated baseline. The baseline is discretionary executive intervention: export controls applied to models, informal pressure before releases, and a voluntary pre-release access framework completed in August 2026 whose text is unpublished and whose coverage thresholds are classified.[5] That is not "less" regulation than a statutory regime. It is regulation without the properties that make regulation legitimate, namely published criteria and a route to challenge a determination. The case for a structured arrangement therefore does not depend on any estimate of catastrophic risk, and the most substantial critiques do not contest the institutional form either, treating independence under industry funding, guaranteed access and the immaturity of measurement as the open questions.[6]

The debate leaves out the record of comparable arrangements, which is extensive and largely public, and the question of who can produce which kind of knowledge, which every proposal answers by implication without asking. Supplying both is the purpose of this paper: sections 2 and 3 ask under which conditions such arrangements work and what the closest comparable regime has produced, section 4 sets out the epistemic division of labour, and section 5 asks what makes the result of a model assessment usable in another regulatory order.

The evidence is comparative, and the closest case is the independent audit regime under Article 37 of the Digital Services Act: two completed cycles of provider-funded external assessment in substantially the same industry, with the reports public.

[4] Google, "A Pragmatic Approach to AI Governance in America" (June 2026), proposing an industry-funded, federally overseen Frontier AI Regulatory Organization on the model of, among others, FINRA and the North American Electric Reliability Organisation; Anthropic, "Anthropic's Advanced AI Framework" (June 2026); OpenAI, "Democratic Governance of Frontier AI: A Blueprint for a Federal Framework" (3 June 2026); M. Thomas, "Designing a FINRA for Frontier AI", Lawfare (30 July 2026).
[5] Executive Order 14409, Promoting Advanced Artificial Intelligence Innovation and Security (2 June 2026), 91 Federal Register 34567, p. 3; completion of the mandated framework reported for 1 August 2026, see M. Curi, "White House finalizes AI framework behind closed doors", Axios (3 August 2026) https://www.axios.com/2026/08/03/white-house-finalizes-ai-framework-behind-closed-doors accessed 16 September 2026.
[6] V. Nguyen, E. Tabassi and K. Duffy, "The U.S. Is About to Design an AI Regulator", Council on Foreign Relations, 23 July 2026; see also, predating the proposals, M. Brundage et al., "Frontier AI Auditing" (2026), arXiv:2601.11699.

2 RESEARCH ON CO-REGULATION

2.1 Concept and state of research

Two terms are in use for the same family of arrangements. Regulated self-regulation is one continental formulation and puts the weight on the private system that a statutory frame takes up; co-regulation is the term of Union practice and describes the combination of non-state and state regulation in such a way that the private system is linked to the public one.[7] Both describe a two-level construction: a statutory framework, and a public body that influences the private process by rule-making and by decisions in individual cases.

Where the private element sits varies. It may lie in rule-setting, where an industry body writes the code a statute takes up, or in enforcement, where a private body applies rules the legislature has set, and the two must be examined separately because a design may privatise one and keep the other public. Subject matters divide into three classes: those where delegation is excluded, those where the legislature may choose, and those closed to state regulation altogether.[8] A further function arises in every variant: someone has to check that what is required is actually being done. Where the public side cannot perform that check, for want of knowledge, access or capacity, or should not, because the proximity it requires would compromise the decision the supervisor later has to take, it passes to a private assessor. That is the audit layer, and section 3 is about what it does.

Measured against these distinctions the American proposals differ. The industry-funded body under federal supervision is co-regulation in the strict sense, a private body exercising regulatory functions on a public legal basis. The framework built on a developer's own commitments is enforced self-regulation with a supervisor attached. And the designated federal agency with mandatory third-party testing is public regulation with the compliance check delegated to private assessors, which is structurally the arrangement of the Digital Services Act. What they have in common, and what this paper takes as its unit of analysis, is the delegation of that check.

The media and communications sector has a long run of such arrangements and a large body of case-study evidence. It has been assembled in studies commissioned by governments and by the European Commission, one of which fed into the co-regulation provisions of Union audiovisual law, and in the research on internet self- and co-regulation.[9] A recent survey of that literature adds a finding that bears on the present case: as platformisation aggravated the regulatory problems, confidence in industry self-regulation declined and initiatives indicating greater state involvement increased.[10]

---

[7] W. Schulz and T. Held, Regulated Self-Regulation as a Form of Modern Government: An Analysis of Case Studies from Media and Telecommunications Law (Eastleigh 2004), 4–8, 16, on the concept and its two-level structure; for the Union formulation, Hans-Bredow-Institut and Institute of European Media Law, Study on Co-Regulation Measures in the Media Sector, Final Report for the European Commission (2006). Not all authors treat the two as interchangeable, and the definitions in use differ. On variants, C. Palzer, "European Provisions for the Establishment of Co-Regulation Frameworks", in European Audiovisual Observatory, IRIS Special: Co-Regulation of the Media in Europe (Strasbourg 2003), 3, 4 f.; W. Hoffmann-Riem, Regulating Media (New York 1996), 326; J. Black, "Constitutionalising Self-Regulation" (1996) 59 MLR 24.

[8] Schulz and Held (n 6), ch. 1, on the distinction between rule-setting and enforcement, on the three classes of subject matter and on barriers to entry, and on the design areas, including time limits, evaluation and provision for failure.

[9] Schulz and Held (n 6); the 2006 study (n 6), prepared for the revision through Directive 2007/65/EC (Art. 4(7) of Directive 2010/13/EU, now Art. 4a as amended by Directive (EU) 2018/1808); J. Cave, C. Marsden and others. Simmons, Options for and Effectiveness of Internet Self- and Co-Regulation (RAND Europe for the European Commission, 2008); C. T. Marsden, Internet Co-Regulation (Cambridge 2011); D. Tambini, D. Leonardi and C. Marsden, Codifying Cyberspace (London 2008).

[10] F. Saurwein, A. Birrer and D. He, "Industry-level Self- and Co-regulation in Media and Communications", in M. Puppis, R. Mansell and H. Van den Bulck (eds), Handbook of Media and Communication Governance (Cheltenham 2024), 245; on private ordering alongside them, T. Mast, M. C. Kettemann and W. Schulz, "Private Ordering of Media Organisations and Platform Operators", ibid., 260. On the oversight board as a reference point for the current debate, K. Yoshino and R. Lemos, "What AI Companies Can Learn from the Oversight Board", Tech Policy Press, 7 July 2026, and K. Frazier, K. Klonick and K. Yoshino, "Scaling Laws: Is Meta's Oversight Board a

Pure self-regulation can take elaborate forms of its own. The oversight board Meta created is independent in composition and publishes reasoned decisions, but rests on a charter rather than a statute and is funded by the company whose decisions it reviews, and it is now being discussed as a model for AI governance. It shows how far publicity and reason-giving can be carried without a public legal basis, which is one of the things the conditions below are meant to measure. The regulator-side assessment frameworks rest on the same base.[11] Two caveats attach to that use: generalisation from case studies is difficult, so what follows are regularities observed in one sector rather than law-like generalisations, and frontier AI differs from media in market concentration, product cycle and the nature of the risks, so each condition is transferred separately and with reasons given.

2.2 Conditions of functioning

Condition 1: A determinate division of labour and responsibility

The condition most consistently identified is that the responsibilities of the public and private participants be set out clearly: the requirement of separation follows from the democratic principle and from the transparency the rule of law demands, and it should be apparent who decides what. Co-regulation does not mean that responsibility for the rules being implemented is shared; the primacy of the public authority remains intact.[12] Where the public side can determine the actions of the private body, the arrangement has stopped being co-regulation and has become a state framework using private expertise, which is a different thing with different legitimation requirements. The reform of the German youth media protection regime, in force since 2003, is the clearest positive model: the recognised self-regulatory bodies classify, and the public commission may overturn a decision only where a body exceeded the limits of professional assessment.[13]

The frontier proposals allocate authority in some detail but leave residual responsibility unallocated. Section 3 discusses what that produces. Where the line runs should follow where each side has its advantage. The reason to use self-regulatory elements at all is that the actors in the field hold the knowledge and see developments first, while the steering capacity of the state declines with a knowledge deficit that grows as the field accelerates; the reason not to leave the field to them is that only the public side can fix the objective and compel production.[14] The distance between supervisor and self-regulatory body is a further design variable with costs on both sides: where the supervisor sits in the drafting groups it is afterwards hard for it to refuse registration, while at a greater distance it loses the information the process generates. [15]

Model for AI Governance?", Lawfare, 27 August 2026.

[11] European Governance: A White Paper, 25 July 2001, COM(2001) 428 final, 28; Mandelkern Group on Better Regulation, Final Report, 13 November 2001, 15 f.; on the regulator side, n 16 below.

[12] Mandelkern Report (n 10), 15.

[13] On the German reform, Schulz and Held (n 6), ch. 5; also ss. 19(2) to (4) and 19a(1) of the Jugendmedienschutz-Staatsvertrag (JMStV) on the recognition and tasks of the self-regulatory bodies, as well as ss. 19b(1) and 20(3) and (5) JMStV, confining the Commission for the Protection of Minors in the Media (KJM) to decisions that exceed the legal limits of the bodies' margin of appreciation (Beurteilungsspielraum).

[14] Schulz and Held (n 6), ch. 1, on the reasons for the failure of traditional regulatory concepts and on the alternatives to them; Saurwein, Birrer and He (n 9).

[15] Schulz and Held (n 6), ch. 2, on proximity and distance between the supervisory authority and the self-regulatory body.

Condition 2: Complexity that remains manageable

Co-regulation multiplies the actors involved, the interfaces between them and the documents crossing those interfaces, and every reform that strengthens one condition adds an interface. Two consequences follow, and both argue for explicit design rather than against co-regulation. Interfaces consume resources that produce no assessment: in medical device conformity assessment, raising assessor quality without expanding supply turned a capture problem into a throughput problem.[16] More dangerous is that responsibility diffuses invisibly. Where several actors each hold part of a function, each can locate the residue with another and none is accountable for the gap; that diffused responsibility along a value chain jeopardises such arrangements unless it is reinforced by direct regulation is a regulator's own finding, drawn from a scheme that failed for that reason.[17] The test is simple: for each function, name the actor who bears it when every other actor declines.

Condition 3: A genuine interest of the regulated actors in their own regulation

The condition recurs across the literature in slightly different words: a common industry interest, meaning a collective will to address the problem, and incentives to participate and comply, of which the credible threat of state intervention is one. In the formulation of one regulator-side assessment, where a substantial gap exists between the public and the private interest it is inappropriate to rely on industry to act in the public interest unless external pressure is applied. The evidence from the media sector adds two qualifications. Advertising self-regulation worked because the industry understood that confidence in advertising as such would suffer if misleading advertising went unchecked, which is an interest in the standard being met, not merely in the existing standard. And where no shared sense of purpose exists, the altruistic premises of co-regulation might not prevail over the narrow commercial interests of the participants.[18]

In the frontier case the interest is real but narrower than it appears. It is an interest in an arrangement, arising from the wish to avoid an unpublished executive gate and a divergent state-level patchwork and from the legitimation a credential confers. It is not an interest in adverse findings about one's own models, and the enlightened self-interest on which such schemes depend is vulnerable to changes in personnel and market structure.

Condition 4: No market participant with divergent interests and the power to act on them

This is the condition the proposals are furthest from satisfying, and the media and telecommunications evidence is unusually direct about it. In the Australian case the same model that worked for broadcasting content failed for interconnection, and the experts attributed the failure to one fact: the former monopolist and its competitors did not have the same interests, and the access forum never performed the functions the statute gave it.[19] Congruent interests are not a soft precondition but the one whose absence has been observed to break the arrangement. The condition also cuts in two directions. A small number of participants with wide coverage facilitates agreement on common rules, which the frontier market has, yet co-regulation is less effective where competition is limited or one large player commands market power

[16] Cf. ErwGr 4 und 9 der VO (EU) 2023/607, Regulation (EU) 2023/607, recitals 4 and 9; on the liability of the notified body, CJEU, judgment of 16 February 2017, C-219/15, Schmitt.

[17] Australian Communications and Media Authority, Optimal Conditions for Effective Self- and Co-Regulatory Arrangements, Occasional Paper (September 2011), 11 ff. and 26; N. Gunningham and J. Rees, "Industry Self-Regulation" (1997) 19 Law & Policy 363; I. Ayres and J. Braithwaite, Responsive Regulation (Oxford 1992), 81 ff.

[18] T. McGonagle, "The Potential for Practice of an Intangible Idea", in IRIS Special (n 6), 15, 23 (advertising) and 24 (absence of a shared sense of purpose).

[19] Schulz and Held (n 6), ch. 2, on the access forum and on interconnection.

the rest of the industry cannot offset.[20] Both features are present at once. Coverage is also incomplete by construction, since pre-release review cannot reach models whose weights are published, and a scheme binding the participants but not the market carries the standing risk that private regulatory organisation produces arrangements operating as cartels.

Agentic deployment sharpens this. The capability is concentrated in a few providers, but agents built on it can be deployed by anyone, so responsibility for what a system does is spread across a chain that a pre-release check does not see. Concentration and diffusion are present at the same time, which is an unusual combination and makes the allocation of responsibility under the first condition even harder.[21]

The risk is not only that such a scheme works badly but that it is unlawful. A body in which competitors agree the standard governing market access is an agreement between undertakings, and the self-regulatory part has to be built to survive that description: open participation and a transparent procedure. Anchoring in public law does not by itself remove the exposure. Under American law a body controlled by market participants is immune only where the state actively supervises it.[22] Funding the body through an assessment on compute providers rather than directly by developers, which is part of the discussion, dilutes the leverage of any single developer and reaches deployments that are not members. It is the most promising element on the table and it does not answer the power question, because leverage in these arrangements is exercised through the choice of assessor and the definition of scope rather than through the invoice.

Condition 5: Residual public authority that is credible

Residual authority has to work without the assessed party's cooperation, and its use has to be plausible. It need not actually be exercised; what matters is that nobody can reasonably assume it never will be. A graduated ladder of sanctions is one way of achieving that, but not the only one. Both American models undercut themselves on this point. The North American Electric Reliability Corporation, named as a template in the proposals, illustrates the problem: its federal supervisor can require a standard to be proposed but cannot attach any consequence to an inadequate response, so its last word carries no weight. How well the federal supervision of self-regulatory organisations has worked is, unusually, a question with a public answer, because a statute requires the supervisor itself to be reviewed at regular intervals. Those reviews found supervision directed at procedures rather than outcomes, and coverage determined by the resources available rather than by risk.[23]

[20] ACMA (n 17), 12 (conditions 1 and 2) and 5 (unintended monopoly power and barriers to entry).
[21] McGonagle (n 20), 18 n. 22.
[22] Arts. 101 and 102 TFEU; Commission Guidelines on horizontal co-operation agreements (2023/C 259/01), on standardisation agreements; CJEU, Case C-124/21 P, International Skating Union (2023). For the United States, Allied Tube & Conduit Corp. v. Indian Head, Inc., 486 U.S. 492 (1988); North Carolina State Board of Dental Examiners v. FTC, 574 U.S. 494 (2015); National Society of Professional Engineers v. United States, 435 U.S. 679 (1978).
[23] U,S, Government Accountability Office, GAO-12-625 (May 2012), GAO-15-376 (April 2015), GAO-18-522 (July 2018), GAO-22-105367 (December 2021), GAO-25-107723 (November 2024), under s. 964 of the Dodd-Frank Act; on the electric reliability model, 16 U.S.C. § 824o.

Condition 6: Observation by third parties, and in public

The enforcement literature adds a third party to the two-player game for a structural reason: where only regulator and regulated are in the room, the regulator is liable to be captured, and outsiders change the game.[24] The media evidence points the same way from the other side: a finding of the Australian study was how little information the arrangement produced for anyone outside it, and one participant took the view that the scheme existed in part to keep the public out of it.[25] Publicity is therefore not an addition to the other conditions but the means by which they can be checked.

Any such arrangement has to reflect the structure of fundamental rights protection it operates within.[26] For frontier models the answer differs with the right at stake. A release decision that affects safety cannot be left to those it binds; a model that shapes what people can read and find out argues against giving the state the last word. Neither side can hold a position the rights at stake deny it. Second, the conditions are not met once and for all, since the costs they impose put them under continuous pressure, which is why time limits and evaluation belong in the design from the outset. These conditions also presuppose an object. It is not a technology but a socio-technical ecosystem: providers, deployers, professional users and the practices that grow up around a system all bear on what it does, and each has to enter the design in its own role. A scheme that addresses the model alone addresses only part of it, which is also why the participation gap noted in section 1 matters for the result and not only for legitimacy.[27]

3 AUDITS UNDER ARTICLE 37 OF THE EU DIGITAL SERVICES ACT

The compliance check described in section 2.1 is not a thought experiment. In one field it already operates under a statute and has run for two cycles with public results, and since 31 August 2026 it reaches frontier providers directly: the Commission has designated ChatGPT a very large online search engine, and the generative answers Google places in its search results ("AI Overviews") sit inside the perimeter of a service designated long ago.[28] Since that designation the same service also falls under the exclusive competence of the AI Office, so the audit under the Digital Services Act and supervision under the AI Act now meet on one product.

Article 37 of the EU Digital Services Act requires providers of very large online platforms and search engines to undergo an annual independent audit of their due diligence obligations, at the provider's expense, by an auditor the provider selects.[29] Whether an audit regime of this kind is itself co-regulation is a question of definition; on the one used in section 2 it is not: the obligations are statutory and the supervisor is public, and what is delegated is the compliance check, to an assessor selected and paid by the assessed party. The same holds for the American proposals: on that definition only the industry-funded body supervised by a federal agency is co-regulation, while the designated-agency model delegates the compliance check in exactly this way.

---

[24] Ayres and Braithwaite (n 17), 54 ff., on third parties as a safeguard against capture; Schulz and Held (n 6), ch. 1.
[25] Schulz and Held (n 6), ch. 2.
[26] White Paper (n 10), 20.
[27] Schulz and Held (n 6), ch. 1, on the design areas.
[28] Commission designation of 31 August 2026 under Art. 33 of Regulation (EU) 2022/2065, see see Commission, "Commission Designates ChatGPT, Reddit, Roblox under Digital Services Act", 31 August 2026; https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1772; accessed 16 September 2026, on some 159 million average monthly recipients of ChatGPT search in the Union, see OpenAI, "EU Digital Services Act (DSA)", OpenAI Help Center; https://help.openai.com/en/articles/8959649-eu-digital-services-act-dsa; accessed 16 September 2026; the additional obligations, the Art. 37 audit among them, apply four months after notification. Google Search has been designated since 2023, so the generative answers presented within it fall inside the audited perimeter; that inference is the author's.
[29] Regulation (EU) 2022/2065, Art. 37(1) and (3); Delegated Regulation (EU) 2024/436, Art. 4.

The record of the DSA regime is therefore instructive about audits rather than about the allocation of regulatory functions, and it is the closest evidence available, being the arrangement now proposed for frontier models applied to the same industry a generation earlier. The resemblance is not merely structural. For the transitional period the most detailed proposal adopts the same design: audits commissioned and paid for by the laboratory, conducted against its own published framework, procedural rather than substantive, reported confidentially to the new body, with remediation before the report is finalised. The first cycle is documented: nineteen reports published at the end of 2024, most by large accounting firms. Auditors were reluctant to interpret the law, restricted themselves largely to reviewing the providers' procedures rather than the sufficiency of their conclusions, and left the systemic risk obligations, which are themselves self-assessment duties, essentially unexamined on substance.[30] Obligations excluded from scope averaged over a hundred sub-articles per service.[31]

One service was audited in both cycles, by firms from different professional traditions. Both audits ended in a negative opinion, but for different reasons: the first because of substantive findings, the second only because obligations had been taken out of scope while the Commission was investigating them. A reciprocity clause attaching to that label would therefore treat an indictment and a scope reduction alike.[32]

Selection and payment create an incentive that professional independence rules have to work against: the assessor is chosen and paid by the party it assesses, and the choice is repeated each year. One of the proposals for AI models anticipates the difficulty and would pool the funding and assign evaluators at random; the others leave it in place. Interpretive authority migrates to the assessed party, which supplied the applicable obligations, their interpretation and the benchmarks; a frontier standard turning on whether a model meaningfully facilitates a cyberattack is at least as open-textured. And private assessment goes dark where public enforcement is live, which is where responsibility visibly diffuses: the auditor withheld judgement because the supervisor was investigating, the provider declined to act for the same reason, and nobody assessed the obligation. Reliance was then foreclosed by the engagement letter behind the report, a statutorily mandated transparency instrument arriving addressed to three named recipients and to nobody else.

What counts as an adequate audit has to be settled by whoever sets up the scheme, not left to the assessor or the party being assessed, since otherwise no two assessments are comparable. A procedural audit also does not answer the question a frontier regime has to answer. Whether a provider ran a risk assessment is not whether its model is dangerous, and in this field the first tells one very little about the second.

[30] D. Holznagel, "Shortcomings of the First DSA Audits and How to Do Better", DSA Observatory, 11 June 2025, on the nineteen reports of 2024 and the procedural treatment of the systemic risk obligations, described there as the auditing of an auditing obligation.
[31] Deloitte, Spotlight on the Year 1 DSA Audit Outcomes" (23 December 2024), reporting an average of 109 of 178 sub-articles ("obligations") being treated as out-of-scope per service; KPMG Netherlands, "Analysis of the 2024 Digital Services Act Audit Reports" (24 February 2025).
[32] FTI Consulting, "X: Independent Audit: Article 37 of Regulation (EU) 2022/2065 (Digital Services Act)", 2023/24; BDO LLP, "X's Digital Services Act (DSA) Report: 24 August 2024 to 30 June 2025", 2024/25.

4 KNOWLEDGE PRODUCTION AND ITS ACTORS

4.1 Actors and their epistemic constraints

Every serious contribution concedes that there is no settled science of frontier capability assessment, and one framework states plainly that a mature independent evaluation ecosystem does not yet exist.[33] The question that follows: Who can produce which kind of knowledge? That regulatory science differs from research science is not new;[34] and neither is their co-evolution, since youth protection supervision developed alongside research on media effects, and platform content moderation alongside the computational social science that studies it. What media regulation knows well is that a measurement discipline and the institution relying on it grow together and slowly. The difference now is that there is no time for that, because the institution is being set up while the discipline that would tell it what to measure is still forming. The differences that matter here are not only differences of interest but also of epistemic constraint, and constraints determine the properties of the product.

Developers alone hold the weights, the training data, the internal-only models, the deployment logs and the ability to run counterfactual experiments during development, and no external arrangement substitutes for this; but a developer cannot credibly report a negative finding about itself, so it contributes access and observation rather than conclusions. Agencies with classified capacity can compare a model against the actual capability of adversaries and can compel production, but their output cannot cumulate publicly, be replicated or enter a foreign process, and a national security frame excludes rights-based effects by construction; they contribute ground truth on threat, the output another supervisor can least rely on. Commercial assurance providers supply methodology, quality management and the capacity to run a recurring programme on a schedule, and section 3 shows what they will not do. Independent research alone produces output that is published, contestable and therefore cumulative.

What makes a finding usable by someone who did not commission it are the properties of academic knowledge production: a published method, a stated measurement uncertainty, a replicable procedure and a record of the method having been contested and survived. A regime composed only of developers, agencies and assurance firms therefore produces findings that are unlikely to be usable outside the issuing state, however rigorous they are internally. Independent research access is usually demanded in the language of accountability and heard as a cost, whereas it is an input to the only process capable of producing findings that others can rely on.[35]

4.2 Functions of independent research

The limits of the research layer are no privileged access, no power to compel, funding neither stable nor timed to releases, and above all a cadence measured in months and years where the object changes in weeks. Peer review, replication and the contestation that gives published findings their standing are slow by design, because that is what makes them reliable. Designs that assign independent research a first-line assessment role therefore fail twice over: the finding arrives after the deployment decision, and the research layer, pressed into a schedule it cannot keep, produces the same procedural output as the assurance layer without the quality apparatus.

The conclusion is not that the layer matters less but that its functions lie one level up, where its cadence is an advantage. Three are available now.

[33] Anthropic (n 3), 7; cf. Nguyen, Tabassi and Duffy (n 5).
[34] S. Jasanoff, The Fifth Branch (Cambridge, Mass. 1990); A. M. Weinberg, "Science and Trans-Science" (1972) 10 Minerva 209.
[35] T. Shevlane, "Structured Access" (2022), arXiv:2201.05159; S. Casper et al., "Black-Box Access Is Insufficient for Rigorous AI Audits" (2024), arXiv:2401.14446.

Function 1: Assessment of the assessment

The first is the assessment of the assessment. The idea is not new: in the enforced self-regulation model the task of the public supervisor is essentially the control of that control, and the case-study literature distinguishes an adequacy check, asking whether the codified rules suffice, from a compliance check, asking whether practice conforms.[36] The only substantive scrutiny the first cycle of Digital Services Act audits received came from independent analysis, which is what established that the reports were procedural, that scope exclusions were systematic and that the systemic risk obligations went rather unexamined.[37] A frontier regime should provide for it rather than depend on it happening.

Function 2: Validation and criticism of measurement instruments

The second is the validation and criticism of measurement instruments, including the unglamorous work of showing which benchmarks saturate, which results carry unreported uncertainty and which claimed capabilities do not replicate. The taxonomy of study designs for investigating platform risks recently produced by a public research institution is a model of the type.[38]

Function 3: Periodic evaluation of the arrangement

The third is periodic evaluation of the arrangement itself against the conditions in section 2.2, on the model of the statutory recurring review that produced the only usable evidence about American self-regulatory supervision. What has to be asked there is institutional rather than technical. Do the interests of the participants still converge, and is the supervisor's residual authority still plausible, and has coverage drifted as new deployers entered the market. None of that can be read off an audit report, and no party to the scheme can answer it without judging its own arrangement. Here the slow cadence is an advantage, provided the evaluation is built in from the start rather than negotiated once the results begin to matter.

All three presuppose the sixth condition of section 2.2, and each presupposes conditions of its own that have to be built into the arrangement rather than left to goodwill.

Access is the scarce good, and only developers and the state can grant it. It should therefore be an obligation of participation, on terms that allow method and results to be published, and matched by publicly funded capacity with compute of its own, because a layer that cannot run the experiments cannot use the access. A classified finding also needs a declassified counterpart, or nobody outside the clearance perimeter can calibrate against it.

Union law already has couplings of this kind, in the vetted researcher access of the Digital Services Act and in the advisory forum and the scientific panel of the AI Act. They were built to feed expertise into enforcement, and adapting them is a smaller step than building something new. Where access is granted only for questions an authority has designated in advance, research turns into an auxiliary of regulation and only what is already thought relevant can be examined.[39]

---

[36] Schulz and Held (n 6), ch. 1, on enforced self-regulation and on the adequacy and compliance checks; Ayres and Braithwaite (n 17), 101 ff.

[37] Holznagel (n 32); that the scope exclusions were systematic rather than particular to one service follows from the benchmark studies at n 31.

[38] C. Panigutti et al., "How to Investigate Algorithmic Driven Risks in Online Platforms and Search Engines? A Narrative Review through the Lens of the EU Digital Services Act", FAccT '25 (2025).

[39] T. Mast and M. Fertmann, "Forschungsdatenzugang und Technologieregulierung: Wissenschaft als Regulierungshilfstätigkeit?"

The most far-reaching version of the access requirement is now proposed from inside the industry: that each frontier company give embedded third-party evaluators (academics might be among those) ongoing, employee-like access, to verify its safety commitments, report incidents and assess training pipelines rather than finished models alone, on the model of embedded banking supervisors. It is the strongest answer to the access and cadence problems on the table, and as a design it concentrates the difficulties above: the evaluator sits closer to the assessed party than any other observer, produces knowledge where publication is least likely and, where cleared, findings that cannot be used abroad. It works only with a determinate allocation of responsibility for acting on what the evaluator sees, and a declassified instrument for what cannot be published.[40]

5 RELIANCE ACROSS REGULATORY ORDERS

5.1 Concept and the position in Union law

What follows takes up a point left open in section 1. The arrangements considered so far are national, the object is not, and the conditions above say nothing about how the results of one might be used in another. The systems are built in one or a handful of jurisdictions and their effects fall everywhere, so results that only their own supervisors can use leave every other jurisdiction deciding without them.

The property at issue is – when we do not consider schemes that are global to start with – whether an authority that did not commission an assessment can take it up, giving it significant weight in its own decision while remaining responsible for that decision. The regulation of medical products has a settled term for this. Reliance is the act by which an authority in one jurisdiction gives significant weight to assessments performed by another authority or trusted institution while remaining independent and accountable for the decision it takes, and it is expressly distinguished from recognition, the routine acceptance of another authority's decision, which dispenses with an assessment of one's own and normally requires binding legal provisions.[41] Interoperability, the term the G7 process employs, names the alignment of governance frameworks as a whole and is the macro-level question.[42] The operative question sits below it, at the level of the individual finding: what the producing side must document and the receiving side must be able to check, which a design can answer, and which leaves the decision with the authority that takes it.

The international level is not short of fora that could help achieve interoperability. The summit series runs from Bletchley Park in 2023 through Seoul and Paris to the India AI Impact Summit of February 2026, and alongside it stand the G7 process, the OECD principles, the network of AI safety institutes, the Global Dialogue on AI Governance and the Independent International Scientific Panel set up by the General Assembly in 2025, the framework convention of the Council of Europe and the legislation of the Union.[43] The deficit is not a missing forum but the weak effect of those that exist and the absence of meaningful

(2024) 57 WissR 101, 121–124, on access regimes that confine research to questions designated in advance and thereby turn it into an auxiliary of regulation.

[40] D. Amodei, "We Must Pace the Frontier" (September 2026), the first of three steps, announced on 12 September 2026 https://darioamodei.com/post/we-must-pace-the-frontier, accessed 16 September 2026; the commitment was announced for OpenAI by S. Altman on X the same day, https://x.com/sama/status/2098811563415150910; accessed 16 September 2026.

[41] WHO, Good Reliance Practices in the Regulation of Medical Products, Annex 10 to WHO Technical Report Series No. 1033 (2021), Glossary and ss. 5.1, 5.2, 5.5; the principles at ss. 6.1 to 6.6 include sovereignty of decision-making and competence.

[42] G7 Digital and Technology Ministers' Statement, Hiroshima AI Process (2023); OECD, Hiroshima AI Process Code of Conduct Reporting Framework, 7 February 2025, both on interoperability between AI governance frameworks.

[43] On the United Nations mechanisms, n 1 above; the non-binding New Delhi Declaration of February 2026 was endorsed by 91 states and organisations; Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (2024).

coupling between them, which is a reliance problem rather than a mandate problem: common terms for security-relevant incidents, a common reporting point and mutual reliance on test results would do more than another body.[44] Union law has made a start, since providers under the competence of the AI Office now report serious incidents to it.

In Union law there is at present nothing for a foreign credential to attach to. The general-purpose AI regime of the AI Act is not a conformity assessment regime in the sense of the New Legislative Framework: there are no notified bodies for models, no third-party certification and no pre-market approval, and compliance is indicated by adherence to a code of practice or, in due course, harmonised standards, supervised directly by the AI Office.[45] The only mechanism for recognising a third-country assessment body sits in the high-risk track and presupposes a Union agreement with that country, and the standardisation route is not available either, since the request in force covers high-risk systems and not models, which also means that whichever body first produces usable benchmarks sets the de facto reference. Recognition is therefore a poor fit for this question, and the experience of data protection suggests that it carries its own risks: the two adequacy decisions for transfers to the United States were annulled precisely at the point where national security intervened, and the challenge to the successor instrument is now before the Court of Justice on appeal[46].

5.2 Ex ante involvement and graded reliance

If recognition after the fact does not carry the weight, the question is how a result can be made usable elsewhere in the first place. The answer that works in practice is to involve those who are later to rely on it before the assessment exists. Under the Medical Device Single Audit Program, one audit by a recognised auditing organisation satisfies the quality system requirements of all participating authorities, which sit as a council and publish the procedures.[47] The media sector supplies a lighter version: the cross-border complaints system of the European advertising self-regulatory bodies handles a complaint in the country of origin under the rules applicable there, on a principle of mutual recognition coordinated through a network secretariat rather than a treaty.[48] What both share is that the authorities relying on the output helped define it, and that reliance comes in degrees. A council-type body in which the AI Office participates in specifying evaluation protocols and recognising evaluators would produce artefacts that office can use, because it will have said what it needs, and this requires no change to Union law.

Where ex ante involvement is not achieved, an evidentiary route remains and requires no instrument at all. The AI Office may request documentation, require model access, conduct its own evaluations and, since the recast of its powers in July 2026, carry out on-site and remote inspections, so if the outputs a frontier body produces are usable in that supervision, de facto reliance follows. That requires comparable taxonomies, documented measurement uncertainty, a disclosure tier open to a foreign supervisor and an allocation of liability permitting reliance, the last being decisive and settled now in engagement letters.

[44] For a common reporting point already in Union law, Regulation (EU) 2024/2847 (Cyber Resilience Act), Art. 14 and 16 (Art. 14 applying since 11 September 2026); further Directive (EU) 2022/2555 (NIS), Art. 23 (incident reporting) and Art. 12 (coordinated vulnerability disclosure); on coordinated vulnerability disclosure, Directive (EU) 2022/2555, Art. 12 and recital 58, referring to ISO/IEC 29147:2018 and ISO/IEC 30111:2019; on the hybrid supervisory architecture this creates, F. Teichmann and B. S. Sergi, "The EU Cyber Resilience Act: Hybrid Governance, Compliance, and Cybersecurity Regulation in the Digital Ecosystem"; (2025) 59 CLSR 106209, 3–5 and 7 f.

[45] Regulation (EU) 2024/1689, Arts. 39 and 88 to 94; Commission Implementing Decision C(2025) 3871 final of 23 June 2025 on a standardisation request as regards high-risk AI systems, repealing C(2023) 3215. The supervisory powers referred to here were extended by Regulation (EU) 2026/1744, in force 27 July 2026, which recast Art. 75 and inserted Arts. 75a to 75d of Regulation (EU) 2024/1689; Chapter IX applies from 2 August 2026.

[46] Case C-362/14 Schrems, judgment of 6 October 2015, and Case C-311/18 Schrems II, judgment of 16 July 2020; on the successor instrument, Commission Implementing Decision (EU) 2023/1795, upheld in Case T-553/23 Latombe, judgment of 3 September 2025, under appeal as Case C-703/25 P.

[47] Medical Device Single Audit Program, Statement of Cooperation and Regulatory Authority Council documents.

[48] McGonagle (n 20), 23 f., on the cross-border complaints system of the European Advertising Standards Alliance, the country-of-origin principle and mutual recognition among the participating bodies.

Product liability and administrative fines work the same way, since neither makes a foreign certification a defence but both make it evidence, whose weight depends on the assessment being legible and attributable [49]

5.3 Jurisdictions with limited evaluation capacity

The transatlantic case is comparatively easy, since the Union has an operating supervisor and the standing to negotiate. For most of the world none of that holds, and the alternative to acceptance is then not abstention: where a jurisdiction has no evaluation capacity of its own, the practical alternative to accepting a certified model is a capable model obtained by another route, including openly released weights, so a credential it can neither verify nor insist upon will not be resisted but ignored. Reliance is a spectrum whose tiers already exist in the comparators above: (i) substitution for the jurisdiction's own procedure, (ii) acceptance in place of routine inspection, (iii) use as an input into a retained domestic decision, and (iv) observer status with capacity building. A regime designed with only the first tier in mind will be adopted by almost nobody, whereas the third and fourth are within reach of many states and are the tiers from which they move upward. Reliance is also asymmetric unless deliberately made mutual, a distinction the medical products guidance draws expressly: a regime in which findings travel outward and nothing inward is not a standard but a channel, and will be understood as one.

Which jurisdictions could occupy which tier is an empirical question nobody has answered, though the instruments exist. Institutional capacity is already mapped across many states by the UNESCO readiness methodology, regional mechanisms provide intermediate addressees where individual states lack capacity, and the capacity to apply that frame country by country sits in the international network of internet and society research centres, the layers of section 4 and absent from every coalition sketch circulating in Washington.[50] The proposal is therefore concrete: a reliance assessment, conducted independently of the body being designed, classifying jurisdictions by the tier they could sustain and identifying what moving up one would require. The caution from the media experience is not merely about texts. When Malaysia adopted an Australian model, the statute travelled and the arrangement did not: two years after entry into force two industry fora had been recognised, the codes had in the end to be drafted by the regulator, and the fora that formed were captured by one part of the industry, which set entry conditions and fees that kept the others out.[51] Transfer requires an implementation concept and a regulator able to moderate it, which is what a reliance assessment measures.

[49] Directive (EU) 2024/2853, Art. 7(1) and (2); for general purpose AI-model providers Regulation (EU) 2024/1689, Art. 101(1); for AI systems Regulation (EU) 2024/1689, Art. 99(7)(e) to (g). Applying the defectiveness test to a third-country certification is an original argument without supporting authority.
[50] UNESCO, Readiness Assessment Methodology (2023); African Union, Continental Artificial Intelligence Strategy (2024).
[51] Schulz and Held (n 6), ch. 3, on Malaysia.

6 CONCLUSION

Co-regulation of frontier AI is a defensible institutional choice, and the present baseline of discretionary executive intervention is no state of freedom. But the current suggestions satisfy few of the conditions under which such arrangements have worked in other sectors.

This paper has argued three things. Whether such an arrangement works has been studied, and the conditions are known well enough to be applied. Independent research is not an add-on but the only source of knowledge a supervisor who was not in the room can check, which places it at the assessment of the assessment rather than at the assessment itself. And whether a result can be used in another jurisdiction is decided when the assessment is designed, not when someone is asked to accept it.

This is a beginning, and it leaves out much that will bear on any design. International competition sets limits on what a jurisdiction can require of developers who can build elsewhere, and open-weight models put capabilities beyond the reach of any pre-release arrangement altogether.

The next step is agreement among enough of the relevant actors on what is to be assessed and reported, and carried by a small group the others trust; an institution may follow. The internet standards community settled long ago on rough consensus and running code. No swarm of coordinated agents will do this for us; someone has to take it in hand.

7 REFERENCES

African Union (2024): Continental Artificial Intelligence Strategy.

Amodei, Dario (2026): We Must Pace the Frontier. Blogpost, 12 September 2026.

Anthropic (2026): Anthropic's Advanced AI Framework. Policy Paper, June 2026.

Australian Communications and Media Authority (2011): Optimal Conditions for Effective Self- and Co-Regulatory Arrangements. Occasional Paper.

Ayres, Ian / Braithwaite, John (1992): Responsive Regulation. Oxford: Oxford University Press.

Black, Julia (1996): Constitutionalising Self-Regulation. In: The Modern Law Review, 59, 24-55.

Brundage, Miles et al. (2026): Frontier AI Auditing. arXiv:2601.11699.

Casper, Stephen et al. (2024): Black-Box Access Is Insufficient for Rigorous AI Audits. arXiv:2401.14446.

Cave, Jonathan / Marsden, Christopher / Simmons, Steven (2008): Options for and Effectiveness of Internet Self- and Co-Regulation. RAND Europe for the European Commission.

Deloitte (2024): Spotlight on the Year 1 DSA Audit Outcomes. Report, 23 December 2024.

European Commission (2001): European Governance: A White Paper. COM(2001) 428 final.

European Commission (2006): Study on Co-Regulation Measures in the Media Sector. Final Report, prepared by Hans-Bredow-Institut and Institute of European Media Law.

Frazier, Kevin / Klonick, Kate / Yoshino, Kenji (2026): Scaling Laws: Is Meta's Oversight Board a Model for AI Governance? In: Lawfare, 27 August 2026.

Google (2026): A Pragmatic Approach to AI Governance in America. Policy Paper, June 2026.

Gunningham, Neil / Rees, Joseph (1997): Industry Self-Regulation. In: Law & Policy, 19, 363-414.

Hoffmann-Riem, Wolfgang (1996): Regulating Media. New York: Guilford Press.

Holznagel, Daniel (2025): Shortcomings of the First DSA Audits and How to Do Better. In: DSA Observatory, 11 June 2025.

Jasanoff, Sheila (1990): The Fifth Branch. Cambridge, Mass.: Harvard University Press.

KPMG Netherlands (2025): Analysis of the 2024 Digital Services Act Audit Reports. Report, 24 February 2025.

Mandelkern Group on Better Regulation (2001): Final Report.

Mast, Torben (2026): From Standards and Codes to Sociotechnical Ecosystems: Co-Regulation in the EU's Digital Single Market. In: ZGE, 18, 88-110.

Mast, Torben / Fertmann, Martin (2024): Forschungsdatenzugang und Technologieregulierung: Wissenschaft als Regulierungshilfstätigkeit? In: WissR, 57, 101-130.

Mast, Torben / Kettemann, Matthias C. / Schulz, Wolfgang (2024): Private Ordering of Media Organisations and Platform Operators. In: Puppis, Manuel / Mansell, Robin / Van den Bulck, Hilde (eds.): Handbook of Media and Communication Governance. Cheltenham: Edward Elgar, 260-275.

Marsden, Christopher T. (2011): Internet Co-Regulation. Cambridge: Cambridge University Press.

McGonagle, Tarlach (2003): The Potential for Practice of an Intangible Idea. In: European Audiovisual Observatory (ed.): IRIS Special: Co-Regulation of the Media in Europe. Strasbourg, 15-24.

Nguyen, Vivian / Tabassi, Elham / Duffy, Kat (2026): The U.S. Is About to Design an AI Regulator. In: Council on Foreign Relations, 23 July 2026.

OECD (2025): Hiroshima AI Process Code of Conduct Reporting Framework. Report, 7 February 2025.

OpenAI (2026): Democratic Governance of Frontier AI: A Blueprint for a Federal Framework. Policy Paper, 3 June 2026.

Palzer, Christoph (2003): European Provisions for the Establishment of Co-Regulation Frameworks. In: European Audiovisual Observatory (ed.): IRIS Special: Co-Regulation of the Media in Europe. Strasbourg, 3-8.

Panigutti, Cecilia et al. (2025): How to Investigate Algorithmic Driven Risks in Online Platforms and Search Engines? A Narrative Review through the Lens of the EU Digital Services Act. In: FAccT '25.

Saurwein, Florian / Birrer, Andrea / He, Duan (2024): Industry-level Self- and Co-regulation in Media and Communications. In: Puppis, Manuel / Mansell, Robin / Van den Bulck, Hilde (eds.): Handbook of Media and Communication Governance. Cheltenham: Edward Elgar, 245-259.

Schulz, Wolfgang / Held, Thorsten (2004): Regulated Self-Regulation as a Form of Modern Government: An Analysis of Case Studies from Media and Telecommunications Law. Eastleigh: John Libbey.

Shevlane, Toby (2022): Structured Access. arXiv:2201.05159.

Tambini, Damian / Leonardi, Danilo / Marsden, Christopher (2008): Codifying Cyberspace. London: Routledge.

Teichmann, Fabian / Sergi, Bruno S. (2025): The EU Cyber Resilience Act: Hybrid Governance, Compliance, and Cybersecurity Regulation in the Digital Ecosystem. In: Computer Law & Security Review, 59, 106209.

Thomas, Marc (2026): Designing a FINRA for Frontier AI. In: Lawfare, 30 July 2026.

U.S. Government Accountability Office (2012): GAO-12-625.

U.S. Government Accountability Office (2015): GAO-15-376.

U.S. Government Accountability Office (2018): GAO-18-522.

U.S. Government Accountability Office (2021): GAO-22-105367.

U.S. Government Accountability Office (2024): GAO-25-107723.

UNESCO (2023): Readiness Assessment Methodology.

Weinberg, Alvin M. (1972): Science and Trans-Science. In: Minerva, 10, 209-222.

WHO (2021): Good Reliance Practices in the Regulation of Medical Products. Annex 10 to WHO Technical Report Series No. 1033.

Yoshino, Kenji / Lemos, Robert (2026): What AI Companies Can Learn from the Oversight Board. In: Tech Policy Press, 7 July 2026.