

The Magnitsky Model for AI: Localized Legislative Cascades as a Response to AI-Driven Circumvention of International Law

Christopher I. V. Farmer, LL.M. (University of Worcester) · Independent Researcher · civfarmer@gmail.com

Preprint; not yet peer reviewed. Phase II Theory Paper. Version 1.1, October 2026.

Abstract

International law’s enforcement architecture is state-centric and treaty-mediated, and both features are now liabilities. Multilateral treaty-making is too slow to track the development of artificial intelligence and too rigid to be corrected once concluded, while the conduct it must govern is increasingly executed not by states but by private systems that diffuse accountability across a corporate form and an algorithmic one. This paper argues that AI has become the perfected proxy: a deniable, jurisdictionally slippery, self-executing intermediary through which individuals, corporations and states can circumvent international-legal norms while remaining formally unaccountable. Extending the author’s prior work on the unregulated character of economic warfare, it contends that the appropriate response is not a new global instrument but its opposite — localised, targeted national legislation on the model of the Sergei Magnitsky and Global Magnitsky Acts. Personal and entity-level measures with extraterritorial bite do not require the consent of the wrongdoer’s state; they attach to individuals and firms, and they cascade through the global financial and market network — dollar clearing, correspondent banking, compliance de-risking and regulatory emulation — to produce de facto harmonisation faster than any convention. The paper develops the treaty-rigidity problem; theorises AI as an accountability-diffusing “double veil”; explains the cascade mechanics of Magnitsky-style designation; and proposes, as an original contribution, a doctrinal mechanism of **AI-conduct-triggered designation** keyed to responsible control rather than to a completed, attributable human wrong. It tests the proposal against worked scenarios and against the strongest objections — sovereignty, over-reach, fragmentation and due process — and offers honest, rather than merely defensive, responses.

Keywords: economic warfare; Magnitsky Act; targeted sanctions; artificial intelligence governance; extraterritoriality; the Brussels effect; corporate veil; responsibility gap; treaty rigidity; legislative cascade.

Suggested citation: Christopher I. V. Farmer, ‘The Magnitsky Model for AI: Localized Legislative Cascades as a Response to AI-Driven Circumvention of International Law’ (preprint, version 1.1, October 2026) <<https://civfarmer.com>>

mer.github.io/publications/phase2-theory/magnitsky-model-for-ai.html>.
<<https://doi.org/10.5281/zenodo.23462144>>.

Licence: Creative Commons Attribution 4.0 International (CC BY 4.0). Law and facts are stated as at 16 October 2026.

1. Introduction

In *Economic Warfare: The Unregulated War*, I argued that the defining pathology of modern conflict is indirection: the party who wills a wrong no longer commits it, but funds, directs or facilitates an interposed actor who commits it in his stead, and so occupies a zone of legal indeterminacy that an international law built for direct inter-state aggression cannot reach.¹ Private military and security companies, corporate intermediaries in conflict-affected extractive regions, cyber-mercenaries available for hire, and private individuals whose deployable capital now outstrips the budgets of many states all trade on a single structural gap: international law fixes responsibility on states, and the sophisticated wrongdoer has learned to insert a deniable intermediary between himself and the act.² The remedy I advanced there was deliberately two-track — an eventual international charter, complemented by domestic legislation that might, by force of example, harden over time into internationally recognised law. This paper returns to that conclusion in the light of the technology that has, since 2024, made the intermediary problem acute, and it reverses the order of priority. The treaty is not the primary remedy. It is, at best, the downstream sediment of a remedy that must begin at the national level.

The intervening development is the maturation of artificial intelligence into an instrument of conduct rather than merely of analysis. An AI system can now execute a course of action — clear an evasive transaction, generate and target a disinformation campaign, price and route a dual-use export, select an objective — at a speed, scale and opacity that no human proxy could match. Crucially, it does so while adding a second layer of deniability on top of the corporate one. Where a private military contractor gave a state *plausible deniability*, an autonomous or semi-autonomous system offers something stronger: a genuine, and genuinely contestable, gap in the chain of human responsibility. The proxy is no longer merely interposed; it is, in part, non-human, and the law of attribution

¹Christopher I V Farmer, ‘Economic Warfare: The Unregulated War’ (LLM dissertation, University of Worcester 2024) chs 1–2 (on indirect, third-party warfare conducted in the interstices of a state-centric legal order) <<https://doi.org/10.5281/zenodo.23462116>> <<https://civfarmer.github.io/publications/economic-warfare/00-economic-warfare-the-unregulated-war-2024-dissertation.pdf>> accessed 16 October 2026. Cited hereafter as Farmer, *Unregulated War*; chapter and section references are to the dissertation as submitted in January 2024.

²ibid ch 1(3)(1) (PMSCs), ch 1(4) (the individual and the state), ch 2 (multinational corporations in conflict-affected regions), ch 3 (cyber-mercenaries) and ch 7(2) (the “democratisation” of economic-warfare capability among ultra-wealthy individuals, multinational corporations and non-state actors).

— designed around human decision — strains to reach through it.³

This paper’s thesis is that the confluence of these two problems — the slowness and rigidity of treaty-based international law, and the accountability-diffusing character of AI — is not answered by more treaty-making. It is answered by the mechanism that international practice has already, if inadvertently, discovered: the Magnitsky model. A single jurisdiction, acting unilaterally, designates a named individual or entity; freezes such assets as fall within its reach; bars that person from its market and financial system; and publishes the designation. Because the modern global economy is a network with a small number of indispensable nodes — the dollar-clearing system, the major correspondent banks, the two or three markets no serious firm can forgo — a designation in one such jurisdiction cascades outward through rational compliance behaviour into a de facto global exclusion, and is then emulated de jure by peer legislatures. No convention is signed. No wrongdoer’s government consents. And yet a norm is enforced.

The argument proceeds in seven further parts. Part 2 develops the treaty-rigidity problem and explains why the instinct to answer AI with a global instrument is misconceived. Part 3 theorises AI as a “double veil” that compounds the corporate veil of company law with the responsibility gap of autonomous systems. Part 4 sets out the mechanics of Magnitsky-style designation and, more importantly, the reasons it cascades. Part 5 offers this paper’s principal original contribution: a doctrinal mechanism of AI-conduct-triggered designation. Part 6 tests the mechanism against three worked scenarios. Part 7 confronts the strongest counterarguments. Part 8 concludes.

2. The Treaty-Rigidity Problem

The reflexive response to a novel transnational harm is to propose a treaty. The reflex is understandable — international law’s most celebrated achievements, from the Geneva Conventions to the Chemical Weapons Convention, are conventional — but it mistakes the pathology. Two features of multilateral treaty-making make it structurally unsuited to governing a technology that changes on an eighteen-month cycle: it is slow to conclude, and it is rigid once concluded.

2.1 The pacing problem

The slowness is not accidental but definitional. A multilateral instrument of any ambition must reconcile the positions of dozens of sovereigns, each with a veto over its own consent, and must then survive domestic ratification in each. The result is a structural lag that Marchant and colleagues have named the **pacing problem**: the exponential character of technological change set

³On plausible deniability as the strategic function of the interposed proxy, see Farmer, *Unregulated War* (n 1) ch 1(3)(1) (PMSCs and Their Utilization).

against the incremental, consensus-bound pace at which legal systems adapt.⁴ The lag is compounded by the **Collingridge dilemma** — that in the early life of a technology, when regulation would be cheap and easy to impose, its eventual effects cannot be known; and by the time those effects are manifest and the case for regulation is plain, the technology is entrenched and regulation has become expensive, slow and resisted.⁵ AI sits squarely in the dilemma’s jaws. The window in which a treaty could have shaped foundation-model development closed before the treaty-drafters convened; what they can now regulate is already load-bearing infrastructure.

The empirical record confirms the diagnosis. The Council of Europe’s Framework Convention on Artificial Intelligence — the world’s first binding international treaty on AI — took two years of intergovernmental negotiation among forty-six member states, the European Union and eleven non-members merely to adopt, and even then only as a *framework* instrument that obliges parties to “adopt or maintain” appropriate domestic measures rather than prescribing them directly.⁶ Its substantive content was progressively diluted toward the lowest common denominator that consensus could bear, with carve-outs for national security and, in significant measure, for the private sector. This is not a criticism of the drafters, who achieved what was achievable; it is an illustration of the ceiling. A treaty produced by consensus among states with divergent interests in AI supremacy will, of necessity, be a treaty about which those states could already agree — which is to say, a treaty that does little the states were not already willing to do.

2.2 Rigidity and ossification

If slowness afflicts a treaty’s birth, rigidity afflicts its life. The Vienna Convention on the Law of Treaties makes amendment of a multilateral treaty a fresh act of treaty-making: absent a bespoke simplified procedure, an amendment binds only those parties that ratify it, so that a large convention cannot be corrected without either unanimity or fragmentation into overlapping legal regimes.⁷ The consequence is ossification. The paradigm case — and one especially apt given the trajectory of autonomous and space-based systems — is the Outer Space Treaty of 1967. Negotiated with dispatch in an earlier technological moment, it has proved practically unamendable, and its Cold-War text now governs, awkwardly and by strained analogy, a domain of mega-constellations, anti-satellite

⁴Gary E Marchant, Braden R Allenby and Joseph R Herkert (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer 2011).

⁵David Collingridge, *The Social Control of Technology* (Frances Pinter 1980).

⁶Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (adopted 17 May 2024, opened for signature 5 September 2024) CETS No 225, arts 1(2), 4, 5 and 7 (obliging parties to “adopt or maintain” domestic measures). The Convention was the product of two years’ work by the Committee on Artificial Intelligence (CAI), comprising the 46 member states, the EU and 11 non-member states.

⁷Vienna Convention on the Law of Treaties (adopted 23 May 1969, entered into force 27 January 1980) 1155 UNTS 331, art 40 (amendment of multilateral treaties).

weapons and contemplated resource extraction that its drafters could not have imagined.⁸ A treaty is a photograph of the technological and political consensus at the instant of its conclusion. For a slow-moving subject that is a virtue. For AI it is a defect, because the photograph is obsolete before the frame is hung.

2.3 The fragmented soft-law alternative is not enough either

The response of the international community to treaty-slowness has been to govern AI, in the interim, through soft law: the OECD AI Principles, endorsed by more than forty states; the UNESCO Recommendation; a proliferation of national strategies and voluntary codes.⁹ Soft law has genuine virtues — it is fast, revisable and coalition-building — and this paper does not disparage it. But soft law cannot bite. It produces a polycentric patchwork of overlapping, competing and non-binding commitments that a determined bad actor simply routes around, and that offers no remedy at all against the individual or entity that deploys an AI system to circumvent a hard-law norm. The governance space is therefore bifurcated: binding instruments that are too slow and too general, and fast instruments that are not binding. What is missing is a mechanism that is simultaneously fast, unilateral and coercive. That description fits the Magnitsky model, and nothing else in the current repertoire fits it as well.

3. Artificial Intelligence as an Accountability-Diffusing Veil

To see why targeted, personal-level measures are the right instrument, one must first understand precisely how AI circumvents international law. It does so by compounding two distinct veils — one drawn from private law, one from the theory of autonomous systems — into a single, unusually opaque barrier between the wrong and the wrongdoer.

3.1 The corporate veil as the original circumvention device

English company law’s foundational principle, established in *Salomon v A Salomon & Co Ltd*, is that a company is a legal person distinct from its members, so that the acts and liabilities of the company are not, without more, the acts and liabilities of those who own or control it.¹⁰ The veil is a deliberate and

⁸Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (opened for signature 27 January 1967, entered into force 10 October 1967) 610 UNTS 205. The Treaty’s amendment provision (art XV) has never been used to modernise its substantive regime; the analogy is offered as illustrative of ossification, not as a claim about space law’s merits.

⁹OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449, adopted 22 May 2019, amended May 2024), endorsed by more than 40 states; UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (adopted 23 November 2021). On the resulting patchwork of principles and its limits, see Anna Jobin, Marcello Ienca and Effy Vayena, ‘The Global Landscape of AI Ethics Guidelines’ (2019) 1 *Nature Machine Intelligence* 389; Brent Mittelstadt, ‘Principles Alone Cannot Guarantee Ethical AI’ (2019) 1 *Nature Machine Intelligence* 501.

¹⁰*Salomon v A Salomon & Co Ltd* [1897] AC 22 (HL).

valuable fiction; but it is also the original accountability-diffusing device, and courts have long recognised that it can be abused. In *Prest v Petrodel Resources Ltd* the Supreme Court treated “piercing” as a remedy of last resort, and Lord Sumption separated two principles that had long been confused: the **concealment principle**, where a company is interposed to conceal the identity of the real actors and the court merely looks behind it, without piercing the veil at all; and the **evasion principle**, where a company is interposed so that its separate personality will defeat, or frustrate the enforcement of, a legal right or obligation that already binds its controller, which on his analysis is the only case in which the veil may be pierced.¹¹ Concealment and evasion are exactly the functions performed by the interposed structures my earlier work examined — the private military company, and the corporate intermediary sourcing minerals from armed-group-held territory — each an entity placed in the chain to hide the instigator and to defeat the norm that would otherwise bind him.¹² The corporate veil, in short, is how the twentieth century laundered responsibility.

3.2 The algorithmic veil: the responsibility gap

Artificial intelligence adds a second veil that the corporate form never possessed. Where the controller of an autonomous or adaptive system deploys it to produce an outcome, the causal chain from human decision to harmful result is genuinely — not merely rhetorically — attenuated. Matthias named this the **responsibility gap**: as machines learn and act on rules their operators did not author and cannot fully predict, there arises a class of harms for which no human satisfies the traditional conditions of responsibility, because none intended, foresaw or directly caused the specific result.¹³ In the humanitarian-law literature the same phenomenon is discussed as the **accountability gap** of autonomous systems, and it has generated proposals as radical as a distinct regime of “war torts” to capture harms that criminal law, built on individual mens rea, cannot reach.¹⁴ The gap is not science fiction; it is the predictable consequence of interposing a decision-making artefact between a principal and an effect.

The algorithmic veil is reinforced by a problem international law already knew it could not solve: **attribution**. In the cyber context, the difficulty of tracing

¹¹*Prest v Petrodel Resources Ltd* [2013] UKSC 34, [2013] 2 AC 415 [16]–[35] (Lord Sumption: the concealment and evasion principles distinguished at [28], where the concealment principle is described as “legally banal” and as not involving piercing the corporate veil at all; the limited evasion principle stated at [35]); [62] (Lord Neuberger: piercing available only “when all other, more conventional, remedies have proved to be of no assistance”).

¹²Farmer, *Unregulated War* (n 1) ch 1(3)(1) (PMSCs as a means of circumventing international obligations while preserving plausible deniability) and ch 2(2)(a) (multinational corporations accused of sourcing minerals from mines controlled by armed groups in the Democratic Republic of Congo); *Adams v Cape Industries plc* [1990] Ch 433 (CA) (group structures and the limits of veil-piercing).

¹³Andreas Matthias, ‘The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata’ (2004) 6 *Ethics and Information Technology* 175.

¹⁴On the accountability gap of autonomous systems in international humanitarian law, and proposals to bridge it, see Rebecca Crootof, ‘War Torts: Accountability for Autonomous Weapons’ (2016) 164 *University of Pennsylvania Law Review* 1347.

conduct to its author — routing through multiple jurisdictions, the deniability of proxy groups, the evidentiary burden of linking a non-state actor to a sponsoring state — is well documented.¹⁵ The applicable secondary rules make the burden severe. Under article 8 of the International Law Commission’s Articles on State Responsibility, the conduct of a non-state actor is attributable to a state only where the actor is in fact acting on the instructions of, or under the direction or control of, that state in carrying out the conduct,¹⁶ a threshold the International Court of Justice set at “effective control” in *Nicaragua*. The looser “overall control” test that the ICTY Appeals Chamber applied in *Tadić* to classify an armed conflict was later rejected by the ICJ as a test of state responsibility in the *Bosnian Genocide* case, which held to the stricter standard.¹⁷ AI supercharges the attribution problem: an autonomous system can act without a contemporaneous human instruction to point to, and the model weights that determine its behaviour are not a “decision” any tribunal can interrogate.

3.3 The double veil and the perfected proxy

Stack the two veils and the circumvention architecture is complete. A wrongdoer incorporates a company in a permissive jurisdiction (veil one); the company owns or licenses an AI system that executes the proscribed conduct (veil two). If the norm-enforcer reaches for the human, the corporate form answers that the *company* acted; if the enforcer reaches for the company, the responsibility gap answers that the *system* acted, autonomously, in a manner its operators neither specified nor could have prevented. Attribution to any sponsoring state fails the *Nicaragua* threshold, because there was no state instruction — only an entity, a model and an outcome. This is the mechanism of AI-driven circumvention of international law, and it is why the twentieth-century toolkit fails. The AI system is the perfected proxy: cheaper than a PMSC, faster than a human network, more deniable than a shell, and — uniquely — able to interpose a real gap in the human chain of responsibility rather than a merely asserted one.

The strategic implication is decisive. Any remedy that depends on attributing a completed, intentional wrong to a state, or to an identifiable human decision, has already lost, because the double veil is engineered precisely to defeat attribution. A workable remedy must therefore do one of two things: it must either abandon the requirement of attributing conduct to a state, or it must relocate the sanctionable fault from the *outcome* the system produced to the *control* the

¹⁵Marco Roscini, *Cyber Operations and the Use of Force in International Law* (Oxford University Press 2014); and see Farmer, *Unregulated War* (n 1) ch 3 (on the difficulty of tracing routed, proxied cyber-operations to their author).

¹⁶International Law Commission, Articles on Responsibility of States for Internationally Wrongful Acts (2001) art 8 (conduct directed or controlled by a State).

¹⁷*Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)* (Merits) [1986] ICJ Rep 14 [115] (“effective control”); *Prosecutor v Tadić* (Judgment) IT-94-1-A (15 July 1999) [120]–[145] (“overall control”); *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro)* (Judgment) [2007] ICJ Rep 43 [398], [403]–[406] (art 8 as customary law; the “overall control” test rejected as a test of state responsibility).

human failed to exercise. The Magnitsky model does the first. The doctrinal mechanism proposed in Part 5 does the second. Together they route around the double veil instead of trying, and failing, to pierce it.

4. The Magnitsky Model: Mechanics and Why It Cascades

The Magnitsky family of instruments is the most important development in the enforcement of transnational norms in a generation, and it is important precisely because of the features that a treaty lacks. It is unilateral, it targets persons rather than states, and its effects propagate through the market without the consent of anyone but the enacting legislature.

4.1 What the instruments do

The original Sergei Magnitsky Rule of Law Accountability Act of 2012 responded to the death in custody of Sergei Magnitsky, who had uncovered a vast tax fraud; the Global Magnitsky Human Rights Accountability Act of 2016 generalised the model from Russia to the world, and Executive Order 13818 (2017) implemented and broadened it, authorising the Treasury, in consultation with the Secretary of State and the Attorney General, to block the property of persons determined to be responsible for serious human-rights abuse or corruption, and suspending their entry into the United States.¹⁸ Operationally, designation places the target on the Office of Foreign Assets Control’s Specially Designated Nationals list; assets within US jurisdiction are frozen; US persons and firms are prohibited from dealing with the target; and the designation is published, imposing a reputational as well as a financial exclusion.¹⁹ The United Kingdom adopted the model in the Global Human Rights Sanctions Regulations 2020, made under the Sanctions and Anti-Money Laundering Act 2018, and the European Union followed within months with its Global Human Rights Sanctions Regime — quickly dubbed the “European Magnitsky Act.”²⁰ Canada, Australia

¹⁸Sergei Magnitsky Rule of Law Accountability Act of 2012, being title IV of the Russia and Moldova Jackson-Vanik Repeal and Sergei Magnitsky Rule of Law Accountability Act of 2012, Pub L 112-208, 126 Stat 1496 (14 December 2012), s 402 (findings); Global Magnitsky Human Rights Accountability Act, Pub L 114-328, div A, title XII, subtitle F, 130 Stat 2533 (23 December 2016); Executive Order 13818, ‘Blocking the Property of Persons Involved in Serious Human Rights Abuse or Corruption’ (20 December 2017) 82 Fed Reg 60839, ss 1(a)(ii) (determinations by the Secretary of the Treasury “in consultation with the Secretary of State and the Attorney General”; “serious human rights abuse”; “corruption”) and 2 (suspension of entry).

¹⁹Global Magnitsky Sanctions Regulations, 31 CFR pt 583 (reissued 89 Fed Reg 17731 (12 March 2024), as amended), s 583.201 and note 1 (names of blocked persons published in the Federal Register and incorporated into the Office of Foreign Assets Control’s Specially Designated Nationals and Blocked Persons List with the identifier “[GLOMAG]”); Executive Order 13818 (n 18) ss 1, 2 and 4. On blocking, the prohibition of transactions by US persons and visa inadmissibility, see Congressional Research Service, *Human Rights and Anti-Corruption Sanctions: The Global Magnitsky Human Rights Accountability Act* (In Focus IF10576, updated 23 July 2026).

²⁰The Global Human Rights Sanctions Regulations 2020, SI 2020/680, made under the Sanctions and Anti-Money Laundering Act 2018 (UK); Council Regulation (EU) 2020/1998

and others have enacted cognate regimes. Two features are doctrinally salient, and I drew attention to both in my earlier work: these instruments fasten on persons and entities rather than on states, and their reach travels with the designated party across borders instead of halting at a territorial frontier.²¹ The measure binds a named actor wherever he is situated, not the territory or the government under which he shelters.

4.2 Why designation cascades: the network mechanism

The decisive property of the model is that a designation in one sufficiently central jurisdiction does not stay in that jurisdiction. It cascades, through three reinforcing channels.

The first is **market access**. The largest consumer and capital markets cannot be forgone by any firm with global ambitions. To be excluded from the United States or the European Union is not a local inconvenience but a commercial catastrophe, and firms will restructure their conduct worldwide to avoid it — the same logic Bradford identified as the **Brussels effect**, whereby the EU externalises its standards through market power alone, without any treaty and without coercion, because multinationals find it cheaper to adopt the strictest applicable rule globally than to maintain divergent product lines.²² The General Data Protection Regulation’s diffusion from Brazil to California is the paradigm; and the EU AI Act now extends the same extraterritorial logic to AI, applying to providers that place systems on the EU market *irrespective of where they are established*, and even to non-EU providers whose systems’ outputs are used within the Union.²³ A designation regime operates on the same principle, in the register of prohibition rather than standard-setting.

of 7 December 2020 concerning restrictive measures against serious human rights violations and abuses [2020] OJ L410I/1 and Council Decision (CFSP) 2020/1999 of 7 December 2020 concerning restrictive measures against serious human rights violations and abuses [2020] OJ L410I/13 (the EU Global Human Rights Sanctions Regime, the “European Magnitsky Act”), both in force 8 December 2020. The initial UK designations (6 July 2020) targeted 49 individuals and organisations, including 25 Russian nationals involved in the mistreatment and death of Sergei Magnitsky: Foreign and Commonwealth Office, ‘UK Announces First Sanctions under New Global Human Rights Regime’ (6 July 2020) <<https://www.gov.uk/government/news/uk-announces-first-sanctions-under-new-global-human-rights-regime>> accessed 16 October 2026.

²¹Farmer, *Unregulated War* (n 1) ch 4(2) (singling out the model’s focus on persons and entities, its reach beyond the enacting state’s borders, and its capacity to be copied by other states as a template).

²²Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020); Anu Bradford, ‘The Brussels Effect’ (2012) 107 *Northwestern University Law Review* 1 (theorising unilateral regulatory globalisation through market power).

²³Regulation (EU) 2016/679 (General Data Protection Regulation) art 3 (territorial scope); Regulation (EU) 2024/1689 (Artificial Intelligence Act), as amended by Regulation (EU) 2026/1744 (Digital Omnibus on AI, OJ L, 2026/1744, 24.7.2026; in force 27 July 2026), art 2(1)(a) and (c) (applying to providers placing systems on the EU market irrespective of establishment, and to non-EU providers whose outputs are used in the Union) and art 22 (authorised representative for non-EU providers of high-risk systems; applicable from 2 December 2027 to systems listed in Annex III and from 2 August 2028 to those covered by Annex I: Regulation (EU) 2026/1744, recital 40).

The second, and more powerful, channel is **financial network centrality** — what Farrell and Newman term **weaponized interdependence**.²⁴ The global financial system is not a flat network but a hub-and-spoke topology with a handful of indispensable nodes: dollar clearing, a small number of correspondent banks, the messaging layer that routes cross-border payments. A US designation reaches far beyond US persons because any non-US firm that deals with the designated party risks its own designation for having materially assisted or supported it and, with it, the loss of its access to dollar clearing and US correspondent accounts — the lever that formal **secondary sanctions** make explicit in other US programmes. Faced with that threat, foreign banks do not litigate; they **de-risk** — they drop the client, because the alternative is to risk their own exclusion from the dollar system, and if one correspondent bank abandons a client the others follow.²⁵ The cascade is thus not primarily coercive at the level of the foreign firm; it is the aggregated, rational, self-protective behaviour of intermediaries who would rather over-comply than gamble their access to the network’s core. This is why a single national list can produce a global exclusion without a single foreign government’s agreement.

The third channel is **legislative emulation**. Once the United States demonstrated that the model worked, peer legislatures copied it — Canada in 2017; Estonia, Latvia and Lithuania ahead of the European Union; the United Kingdom and the European Union in 2020; Australia in 2021 — converting a de facto cascade into a de jure convergence.²⁶ The origins of the sanctions instrument as a blunt collective weapon are documented by Mulder, and its refinement from the 1990s into a precise tool of individualised statecraft by the targeted-sanctions literature;²⁷ the Magnitsky innovation was to marry that precision to a human-rights trigger and an extraterritorial reach. The point for present purposes is that the diffusion happened *through domestic enactment inspired by*

²⁴Henry Farrell and Abraham L Newman, ‘Weaponized Interdependence: How Global Economic Networks Shape State Coercion’ (2019) 44(1) *International Security* 42.

²⁵Executive Order 13818 (n 18) s 1(a)(iii)(A) (designation of any person determined to have “materially assisted, sponsored, or provided financial, material, or technological support for” a sanctioned activity or a blocked person); on secondary sanctions, which force foreign firms to choose between trading with US sanctions targets and forfeiting access to the US market, Tom Ruys and Cedric Ryngaert, ‘Secondary Sanctions: A Weapon out of Control? The International Legality of, and European Responses to, US Secondary Sanctions’ (2019) 89 *British Yearbook of International Law* 12; and on the leverage that control of network hubs gives the state that hosts them, Farrell and Newman (n 24).

²⁶See nn 18–20; Martin Russell, ‘Global Human Rights Sanctions: Mapping Magnitsky Laws — The US, Canadian, UK and EU Approach’ (European Parliamentary Research Service Briefing PE 698.791, November 2021) <[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698791/EPRS_BRI\(2021\)698791_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698791/EPRS_BRI(2021)698791_EN.pdf)> accessed 16 October 2026 (documenting the diffusion of Magnitsky-style regimes, including Canada’s 2017 Act and the measures adopted by Estonia, Latvia and Lithuania before the EU regime); Autonomous Sanctions Amendment (Magnitsky-style and Other Thematic Sanctions) Act 2021 (Cth).

²⁷Nicholas Mulder, *The Economic Weapon: The Rise of Sanctions as a Tool of Modern War* (Yale University Press 2022); Daniel W Drezner, ‘Sanctions Sometimes Smart: Targeted Sanctions in Theory and Practice’ (2011) 13 *International Studies Review* 96; Thomas J Biersteker, Sue E Eckert and Marcos Tourinho (eds), *Targeted Sanctions: The Impacts and Effectiveness of United Nations Action* (Cambridge University Press 2016).

example — precisely the pathway my earlier work anticipated, in which a purely internal measure is taken up abroad until it operates as a shared international standard — and it happened in years, not the decades a treaty would have consumed.

4.3 The upshot

The Magnitsky model therefore possesses the three properties that Part 2 found treaties to lack and soft law to lack differently: it is fast (a designation can be made in weeks), it is unilateral (it needs no wrongdoer’s consent), and it is coercive (it bites through the market). It does not require attribution of conduct to a state, and it does not require a completed inter-state wrong. It requires only a named person, a sanctionable predicate, and a jurisdiction central enough to make exclusion hurt. That is a far lighter evidentiary and diplomatic load than a treaty imposes — and it is a load that, as the next Part argues, can be adapted to the specific predicate of AI-driven circumvention.

5. A Doctrinal Mechanism: AI-Conduct-Triggered Designation

The existing Magnitsky instruments are keyed to serious human-rights abuse and corruption. They are not, as drafted, apt to reach the person who deploys an AI system to circumvent an international-legal norm, because their triggers presuppose a completed, humanly-authored wrong. This Part proposes — as the author’s original contribution, and offered as a model for legislative adaptation rather than as a statement of existing law — a new designation trigger designed for the double veil: **AI-conduct-triggered designation**.

5.1 The core move: from outcome-attribution to responsible control

The animating insight is the one identified at the close of Part 3. The double veil defeats any remedy that must attribute a proscribed *outcome* to a human intention. The mechanism must therefore relocate the sanctionable fault from the outcome to the **control**: it must make the failure to maintain adequate human control over a system that produces a proscribed outcome *itself* the designable conduct. On this construction the responsibility gap ceases to be a shield and becomes the trigger. A designated person cannot answer “the system acted, not I,” because the gravamen of the designation is not that he acted but that he *deployed a system he did not adequately control* to a proscribed end from which he benefited.

This is not a novel jurisprudential invention so much as the application of settled ideas to a new object. Three established doctrines supply the analogy. First, **superior responsibility** in international criminal law holds a commander liable for the acts of subordinates he failed to prevent or punish, where he knew or should have known and failed to take reasonable measures — a liabil-

ity grounded in defective control, not in the commander’s own commission.²⁸ Second, the **failure-to-prevent** model in domestic economic-crime law — the corporate offence of failing to prevent bribery under section 7 of the Bribery Act 2010, and the failure-to-prevent-fraud offence introduced by section 199 of the Economic Crime and Corporate Transparency Act 2023 — imposes liability on an organisation for a wrong committed by a person associated with it unless it proves that it had in place adequate procedures (the 2010 Act) or such prevention procedures as were reasonable in all the circumstances (the 2023 Act, which applies only to large organisations).²⁹ Third, product-liability logic locates responsibility with the party best placed to control a dangerous instrumentality. Each relocates fault from the completed wrong to the defendant’s relationship of control over the instrument of wrong. AI-conduct-triggered designation does the same, in the register of administrative sanction rather than criminal or tortious liability.

5.2 The proposed elements

A designation under the proposed trigger should require the enacting authority to be satisfied, to a defined evidentiary standard, of three elements.

(i) **A proscribed outcome — the norm-circumvention predicate.** The AI system must have produced, or materially enabled, an outcome that circumvents a specified norm of international law: for example, the evasion of a lawfully imposed sanctions regime; the conduct of a cyber-operation against critical infrastructure of a kind that would, if attributable to a state, breach the prohibition on intervention; the material facilitation of an act that would constitute a war crime; or the manipulation of a market or an electoral process by automated means. The predicate norms should be enumerated, not open-ended, to constrain executive discretion (see 5.4).

(ii) **A responsible-control nexus.** The person to be designated must stand in a defined relationship of responsibility to the system: ownership, effective operational control, deployment for their benefit, or knowing and material provision of the system to the proximate actor. This is the analogue of the concealment/evasion enquiry in *Prest*: the designation reaches the true actor behind the interposed system, whether the interposition is corporate, algorithmic, or both.

²⁸Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3, art 28 (responsibility of commanders and other superiors for failure to prevent or punish).

²⁹Bribery Act 2010 (UK) s 7 (failure of commercial organisations to prevent bribery; defence of “adequate procedures”, s 7(2)); Economic Crime and Corporate Transparency Act 2023 (UK) s 199 (failure to prevent fraud; the offence applies to “a relevant body which is a large organisation”, s 199(1); defence that the body had in place “such prevention procedures as it was reasonable in all the circumstances to expect the body to have in place”, or that it was not reasonable to expect it to have any, s 199(4); in force 1 September 2025: Economic Crime and Corporate Transparency Act 2023 (Commencement No 4) Regulations 2025, SI 2025/349, reg 3).

(iii) **A fault nexus.** The person must have acted with the requisite fault — intention or recklessness as to the proscribed outcome, *or*, in the alternative and critically, a **failure to maintain meaningful human control** and adequate governance over a system whose deployment created a foreseeable risk of that outcome. It is this alternative limb that closes the responsibility gap: the operator who cannot be shown to have intended the outcome may still be designated if he deployed a high-risk system without the human-in-the-loop safeguards, auditability and impact assessment that reasonable governance required. The standard is deliberately modelled on the defences to the failure-to-prevent offences, under which the organisation must prove adequate (2010 Act) or reasonable (2023 Act) prevention procedures: the burden of demonstrating meaningful human control should rest, once the first two elements are established, on the person best placed to know how the system was governed — subject to the safeguards discussed below.

5.3 Evidentiary approach

Two evidentiary features follow from the double-veil diagnosis. First, because algorithmic opacity impedes the designating authority as much as it shields the target, the standard of proof should be the administrative “reasonable grounds to suspect” threshold already familiar from the United Kingdom’s Magnitsky regime, not the criminal standard — designation is a preventive market-exclusion measure, not a conviction. Second, once the authority establishes the proscribed outcome and the control nexus on that standard, a limited evidential burden should shift to the person to establish meaningful human control, because the facts of the system’s governance lie peculiarly within his knowledge. This burden-shift is the mechanism’s most contestable feature and is addressed squarely in Part 7.

5.4 Safeguards designed in, not bolted on

The Magnitsky instruments have been fairly criticised for weak ex ante control and for the risk of politicised or arbitrary listing — concerns about the reach of executive power that I examined in my earlier work, and that were voiced while SAML A was still a Bill, when Lord McNally, supporting an amendment to remove its power to create criminal offences by regulations, told peers that voting for it would “send a message which will avoid a constitutional car crash further down the road”.³⁰ A mechanism that lowers the threshold for designation to a control failure must therefore build its due-process protections into the trigger itself, not leave them to subsequent challenge. The proposal accordingly

³⁰HL Deb 15 January 2018, vol 788, col 483 (Lord McNally, at the Report stage of the Sanctions and Anti-Money Laundering Bill [HL], on Amendment 45, moved by Lord Judge to leave out para (a) of cl 16 and so remove the power for sanctions regulations to create criminal offences; agreed on division, 209 to 192). The remark concerned the Bill’s delegated powers, not its designation criteria. On the breadth of executive discretion under SAML A and the difficulty for targets of securing the lifting of sanctions, see Farmer, *Unregulated War* (n 1) ch 4(1).

incorporates: an **enumerated** and closed list of predicate norms; a published statement of reasons for each designation, to the fullest extent security permits; an **independent review** body with power to recommend delisting; a **delisting pathway** available on demonstration of remediation (divestment, decommissioning or the retrofitting of meaningful human control); and a **sunset and periodic-review** requirement, so that a designation lapses unless affirmatively renewed on current evidence. These are not concessions grudgingly appended; they are constitutive of a model that aspires to be a rule-of-law instrument rather than a mere weapon.

5.5 Why a designation, and not a new offence

It is worth stating plainly why the mechanism is a *designation* — an administrative, extraterritorial market-exclusion — rather than a new criminal offence or a treaty obligation. A criminal offence would run straight back into the attribution and mens rea problems that the double veil is built to exploit, and would be confined to the enacting state’s territorial jurisdiction and its willing extradition partners. A treaty would be slow and rigid for the reasons given in Part 2. A designation, by contrast, needs no attribution to a state, no completed intentional wrong, and no foreign consent; it needs only the enacting jurisdiction’s sovereign prerogative to decide who may access its market and its financial system — and, through the cascade, that single decision does the work of a treaty that no one had to sign.

6. Worked Scenarios

The mechanism is best understood in operation. Three scenarios, deliberately spanning the economic, informational and kinetic registers of the double veil, illustrate both its reach and its limits.

6.1 The algorithmic sanctions-evasion engine

A trading company incorporated in a permissive offshore jurisdiction licenses a proprietary AI system that continuously restructures ownership chains, re-invoices and re-routes shipments, and times transactions to defeat pattern-detection, enabling a sanctioned entity to move dual-use goods through the international financial system. No employee “decides” any particular evasive transaction; the system does, autonomously, thousands of times a day. Traditional enforcement founders: there is no attributable state instruction, the corporate form conceals the beneficiaries, and no human authored the specific evasive act.

Under AI-conduct-triggered designation the analysis is straightforward. The proscribed-outcome element is satisfied by the systematic circumvention of a lawful sanctions regime (an enumerated predicate). The control-nexus element reaches the beneficial owners and the officers who deployed the system for their benefit. The fault element is satisfied either by recklessness — deploying a

system whose designed purpose is evasion — or, failing proof of that, by the manifest absence of meaningful human control over a system engineered to act at machine speed precisely so that no human would review its acts. Designation of the owners and the entity triggers the cascade: correspondent banks de-risk, the offshore vehicle loses access to dollar clearing, counterparties in third states drop it to protect their own network access, and the evasion engine is starved of the financial rails it exists to exploit. No treaty; no attribution to any state; weeks, not decades.

6.2 The autonomous influence operation

A private firm sells, as a service, an AI system that generates and micro-targets synthetic political content at scale to destabilise a foreign election or manipulate a securities market. The client is a non-state actor; the sponsoring state, if any, is undetectable behind the *Nicaragua* threshold. The harm is real but attribution is hopeless, and the content itself may be lawful speech in the vendor’s home jurisdiction.

Here the mechanism’s limits are as instructive as its reach. Designation cannot, and should not, target the *speech*; it targets the *person who deployed an automated manipulation system* to a proscribed end. The predicate must be carefully enumerated — automated manipulation of an electoral or market process, not “disinformation” at large — precisely to avoid the free-expression over-reach that Part 7 concedes is a genuine risk. Where the predicate is met, the vendor firm and its principals are designable on the control-and-fault nexus, and the cascade excludes them from the payment processors, ad exchanges and banking relationships on which a commercial influence-operation depends. The scenario shows that the mechanism is powerful against *commercial* deployers who need market access, and correspondingly weaker against ideologically-motivated actors indifferent to it — an honest limitation, not a defect to be papered over.

6.3 The deniable autonomous-targeting service

A defence-technology company, structured through a chain of subsidiaries, licenses an AI targeting system to a party to an armed conflict; the system selects and prioritises objectives with minimal human review, and its use contributes to strikes that would, if attributable to a state, constitute violations of international humanitarian law. The vendor asserts that it merely supplied software; the operator asserts that the system, not any commander, selected the objective; the sponsoring state asserts nothing, because nothing can be attributed to it.

This is the double veil in its gravest form, and it is where the mechanism most directly answers the accountability-gap literature.³¹ The proscribed outcome is the facilitation of conduct that would breach international humanitarian law. The control nexus reaches both the vendor that knowingly supplied a system for that use and the operator that deployed it without meaningful human control.

³¹See Crootof (n 14) and the accountability-gap literature discussed in Part 3.2.

The fault element is squarely engaged by the *failure-to-control* limb: a targeting system deployed with “minimal human review” is the paradigm of the abdication the limb is designed to sanction. Designation will not, by itself, end the conflict or substitute for the International Criminal Court; but it strips the vendor and the deploying entity of the market and financial access on which even defence contractors depend, and it does so without waiting for an attribution that will never come. The scenario also marks the mechanism’s boundary: against a state actor with a captive domestic market and no need of the dollar system, the cascade weakens — a limit acknowledged, not concealed.

7. Counterarguments and Honest Responses

A proposal of this kind earns its keep only by confronting the strongest objections to it. Four are serious. I state each at its full strength and answer it honestly, conceding what must be conceded.

7.1 Sovereignty and extraterritoriality

The objection. The mechanism is an extraterritorial exercise of power. It reaches persons and conduct beyond the enacting state’s borders, coerces foreign firms through their bankers, and — a difficulty I acknowledged of SAMLA in my earlier work — is liable to be received abroad as an affront to sovereignty, straining diplomatic relations and inviting the charge that a handful of financially central states are legislating for the world.³² The complaint that unilateral sanctions amount to economic bullying, and that their humanitarian costs fall on the innocent, is not met by the retort that the target is a wrongdoer.³³

The response. Two points, one formal and one candid. Formally, a designation is an exercise of a state’s undoubted sovereign prerogative to determine who may access *its own* market and financial system; it prescribes no rule of conduct for foreign territory and imposes no penalty enforceable abroad. The extraterritorial *effect* is produced by the voluntary, self-protective compliance choices of intermediaries, not by any claim of prescriptive jurisdiction over foreign conduct. That is a meaningful distinction between the Magnitsky model and a genuine extraterritorial criminal law. Candidly, however, the distinction is formal, and the coercion is real: the cascade works precisely because exclusion from the network is intolerable, and a mechanism that depends on that intolerance cannot honestly deny that it coerces. The proper response is not to pretend otherwise but to constrain the power — through the enumerated predicates and independent review of Part 5 — and to accept that its legitimacy rests, ultimately, on the legitimacy of the norm it enforces. A cascade in service of a defensible, enumerated norm is statecraft; a cascade in service of a pretext is the abuse the safeguards exist to prevent.

³²Farmer, *Unregulated War* (n 1) ch 4(1) (noting that the extraterritorial reach of SAMLA risks being seen abroad as an incursion on sovereignty and can strain diplomatic relations).

³³Joy Gordon, *Invisible War: The United States and the Iraq Sanctions* (Harvard University Press 2010) (on the humanitarian costs of broad sanctions).

7.2 Over-reach, politicisation and the rule of law

The objection. Designation regimes are executive instruments with weak ex ante judicial control. They can be, and have been, deployed for political convenience; the lowered threshold this paper proposes — designation on a *failure to control*, with a shifted evidential burden — magnifies the danger, sweeping in the merely negligent alongside the culpable and inviting arbitrary or pretextual listing.

The response. This is the objection I take most seriously, because it is partly correct. A mechanism that trades some of the rigour of the criminal process for speed and reach does incur a rule-of-law cost, and it is not honest to deny it. The mitigation is threefold and is built into the trigger rather than left to chance: the predicate norms are closed and enumerated, so the executive cannot designate at large; designations carry published reasons and are subject to independent review and a delisting pathway; and the fault threshold, while it includes a control-failure limb, is not strict liability — the person who can demonstrate meaningful human control is not designable. The evidential burden-shift is confined to the governance facts peculiarly within the target’s knowledge, and operates only after the authority has independently established the proscribed outcome and the control nexus. These constraints do not eliminate the risk; they bound it. The residual risk is the price of a fast instrument, and the paper’s claim is comparative, not absolute: this risk is smaller than the risk of leaving AI-driven circumvention wholly unremedied while a treaty is negotiated over a decade.

7.3 Fragmentation and conflicting designations

The objection. If many states enact divergent triggers and lists, the result is not convergence but a fractured compliance landscape in which a firm is required by one jurisdiction to do what another forbids — the very conflict already visible between US secondary sanctions and the EU Blocking Regulation, which prohibits EU persons from complying with certain foreign measures.³⁴ Fragmentation could be weaponised, each bloc designating the other’s champions, and the global network could balkanise into incompatible spheres.

The response. The risk is real and I do not minimise it, but the cascade dynamics cut against it. De-risking is an *over*-compliance behaviour: when in doubt, intermediaries drop the client, which pushes the system toward the stricter rule rather than toward stalemate. The empirical record of Magnitsky adoption shows convergence, not fragmentation — the UK and EU emulated the US model rather than contradicting it — because the enacting states shared

³⁴Council Regulation (EC) 2271/96 of 22 November 1996 protecting against the effects of the extra-territorial application of legislation adopted by a third country, and actions based thereon or resulting therefrom [1996] OJ L309/1 (the EU Blocking Statute), as amended, art 5 (prohibiting EU persons from complying with the third-country measures specified in its Annex, which was replaced by Commission Delegated Regulation (EU) 2018/1100 [2018] OJ L199I/1).

the underlying norm.³⁵ Fragmentation becomes acute only where the underlying norms genuinely conflict, and there the problem is not the instrument but the absence of normative consensus, which no instrument can manufacture. The honest concession is that AI-conduct-triggered designation works best among states that already share the predicate norms, and degrades precisely where geopolitical rivals do not — which is to say it is a tool for enforcing consensus, not for creating it. Interoperability measures — mutual recognition of designations among aligned states, shared evidentiary standards — can widen the zone of convergence, but cannot abolish the rivalry at its edge.

7.4 Due process, opacity and the innovation chill

The objection. Designating a person for what a machine did, on a shifted burden, in the face of algorithmic opacity that the target may be no better placed than the authority to dispel, risks grave unfairness. And at the system level, a broad control-failure trigger could chill legitimate AI development, pushing research into opaque jurisdictions and deterring exactly the safety-conscious developers the mechanism should encourage.

The response. On due process, the mitigations of Part 5 are directly responsive: the burden-shift is limited to governance facts within the target’s knowledge, the standard is administrative rather than criminal, and independent review with a remediation-based delisting pathway gives the target a route out that a conviction would not. It remains true that opacity can cut against the target as well as the authority, and the mechanism should therefore treat *demonstrable* meaningful human control — auditability, human-in-the-loop design, documented impact assessment — as a complete answer, which converts the trigger into an incentive: the developer who governs his system well is safe, and is *safer* than a competitor who does not. That is the reply to the innovation-chill objection as well. Far from penalising responsible development, a control-failure trigger rewards it, and aligns the private incentive with the public interest in “regulated-by-design” AI. The developers who have something to fear are those who deploy high-risk systems without the governance the trigger rewards — which is the population the mechanism is meant to reach. The residual concern, that determined actors will migrate to jurisdictions beyond the cascade’s reach, is answered only partially and only by the network’s centrality: the migration is available, but the market and financial access the migrant forgoes is the price, and for commercially-motivated actors that price is often prohibitive. For actors indifferent to it, the mechanism is honestly acknowledged to be weaker — as scenario 6.2 showed — and must be supplemented by other tools.

³⁵See nn 20 and 26 (UK and EU emulation of, rather than divergence from, the US model).

8. Conclusion

The argument of this paper is best understood as an evolution of, and a correction to, my earlier position. *Economic Warfare: The Unregulated War* diagnosed the disease correctly — the interposition of deniable proxies to circumvent a state-centric international law — but it hedged on the cure, proposing both an eventual international charter and domestic legislation to inspire it. The intervening rise of artificial intelligence as an instrument of conduct has resolved the hedge. The treaty is not the remedy; it is too slow to be born and too rigid to be corrected, and it would in any event founder on the very attribution problem that the AI-augmented double veil is engineered to exploit. The remedy is the mechanism that the Magnitsky instruments discovered by practice and that this paper has sought to theorise and to extend: localised, targeted, personal-level legislation with extraterritorial bite, which needs no wrongdoer’s consent, attaches to individuals and entities rather than to states, relocates the sanctionable fault from an unprovable outcome to a demonstrable failure of control, and cascades through the market’s indispensable nodes into a de facto global norm.

This is not a claim that designation is a panacea. Part 7 conceded its coerciveness, its rule-of-law costs, its dependence on shared norms, and its weakness against actors indifferent to market access. It is a claim about sequence and comparison. Against a harm that is already operating and a treaty process that will take a decade, the localised legislative cascade is the instrument that can act now, and — this is the deeper point — the norms it enforces, replicated across enacting legislatures and hardened by market practice into settled expectation, are precisely the substrate on which a durable treaty could one day be built. The cascade does not displace the convention; it precedes it, as domestic practice has always preceded and shaped customary and conventional international law. The choice, in other words, is not between the Magnitsky model and a treaty. It is between enforcing the norm now, by cascade, and waiting to enforce it later, by a convention that the technology will have outrun before the ink is dry. For the accountability-diffusing machine, the answer is not a slower, grander machine of international law. It is a smaller, faster, and sharper one — aimed, this time, at the human who hides behind the veil.

Notes

References

Primary sources — legislation, treaties and instruments

- Autonomous Sanctions Amendment (Magnitsky-style and Other Thematic Sanctions) Act 2021 (Cth).
- Bribery Act 2010 (UK).

- Commission Delegated Regulation (EU) 2018/1100 of 6 June 2018 amending the Annex to Council Regulation (EC) No 2271/96 [2018] OJ L199I/1.
- Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (adopted 17 May 2024, opened for signature 5 September 2024) CETS No 225.
- Council Decision (CFSP) 2020/1999 of 7 December 2020 concerning restrictive measures against serious human rights violations and abuses [2020] OJ L410I/13.
- Council Regulation (EC) 2271/96 of 22 November 1996 (EU Blocking Statute) [1996] OJ L309/1, as amended.
- Council Regulation (EU) 2020/1998 of 7 December 2020 concerning restrictive measures against serious human rights violations and abuses (EU Global Human Rights Sanctions Regime) [2020] OJ L410I/1.
- Economic Crime and Corporate Transparency Act 2023 (UK), s 199.
- Economic Crime and Corporate Transparency Act 2023 (Commencement No 4) Regulations 2025, SI 2025/349 (UK).
- Executive Order 13818, ‘Blocking the Property of Persons Involved in Serious Human Rights Abuse or Corruption’ (20 December 2017) 82 Fed Reg 60839.
- Global Magnitsky Human Rights Accountability Act, Pub L 114-328, div A, title XII, subtitle F, 130 Stat 2533 (2016).
- Global Magnitsky Sanctions Regulations, 31 CFR pt 583 (reissued 89 Fed Reg 17731 (12 March 2024), as amended).
- International Law Commission, Articles on Responsibility of States for Internationally Wrongful Acts (2001).
- OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449, 2019, amended 2024).
- Regulation (EU) 2016/679 (General Data Protection Regulation).
- Regulation (EU) 2024/1689 (Artificial Intelligence Act), as amended by Regulation (EU) 2026/1744.
- Regulation (EU) 2026/1744 of 8 July 2026 (Digital Omnibus on AI) OJ L, 2026/1744, 24.7.2026.
- Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3.
- Sanctions and Anti-Money Laundering Act 2018 (UK).
- Sergei Magnitsky Rule of Law Accountability Act of 2012, Pub L 112-208, title IV, 126 Stat 1496.
- The Global Human Rights Sanctions Regulations 2020, SI 2020/680 (UK).
- Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space (opened for signature 27 January 1967, entered into force 10 October 1967) 610 UNTS 205.
- UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (2021).
- Vienna Convention on the Law of Treaties (adopted 23 May 1969, entered into force 27 January 1980) 1155 UNTS 331.

Primary sources — cases

- *Adams v Cape Industries plc* [1990] Ch 433 (CA).
- *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro)* (Judgment) [2007] ICJ Rep 43.
- *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)* (Merits) [1986] ICJ Rep 14.
- *Prest v Petrodel Resources Ltd* [2013] UKSC 34, [2013] 2 AC 415.
- *Prosecutor v Tadić* (Judgment) IT-94-1-A (15 July 1999).
- *Salomon v A Salomon & Co Ltd* [1897] AC 22 (HL).

Primary sources — parliamentary debates

- HL Deb 15 January 2018, vol 788, col 483.

Secondary sources — books, articles and reports

- Biersteker T J, Eckert S E and Tourinho M (eds), *Targeted Sanctions: The Impacts and Effectiveness of United Nations Action* (Cambridge University Press 2016).
- Bradford A, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020).
- Bradford A, ‘The Brussels Effect’ (2012) 107 Northwestern University Law Review 1.
- Collingridge D, *The Social Control of Technology* (Frances Pinter 1980).
- Congressional Research Service, *Human Rights and Anti-Corruption Sanctions: The Global Magnitsky Human Rights Accountability Act* (In Focus IF10576, updated 23 July 2026).
- Crootoof R, ‘War Torts: Accountability for Autonomous Weapons’ (2016) 164 University of Pennsylvania Law Review 1347.
- Drezner D W, ‘Sanctions Sometimes Smart: Targeted Sanctions in Theory and Practice’ (2011) 13 International Studies Review 96.
- Farmer C I V, ‘Economic Warfare: The Unregulated War’ (LLM dissertation, University of Worcester 2024).
- Farrell H and Newman A L, ‘Weaponized Interdependence: How Global Economic Networks Shape State Coercion’ (2019) 44(1) International Security 42.
- Foreign and Commonwealth Office, ‘UK Announces First Sanctions under New Global Human Rights Regime’ (6 July 2020).
- Gordon J, *Invisible War: The United States and the Iraq Sanctions* (Harvard University Press 2010).
- Hufbauer G C, Schott J J, Elliott K A and Oegg B, *Economic Sanctions Reconsidered* (3rd edn, Peterson Institute for International Economics 2007).
- Jobin A, Ienca M and Vayena E, ‘The Global Landscape of AI Ethics Guidelines’ (2019) 1 Nature Machine Intelligence 389.

- Marchant G E, Allenby B R and Herkert J R (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer 2011).
- Matthias A, ‘The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata’ (2004) 6 *Ethics and Information Technology* 175.
- Mittelstadt B, ‘Principles Alone Cannot Guarantee Ethical AI’ (2019) 1 *Nature Machine Intelligence* 501.
- Mulder N, *The Economic Weapon: The Rise of Sanctions as a Tool of Modern War* (Yale University Press 2022).
- National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1, 2023).
- Roscini M, *Cyber Operations and the Use of Force in International Law* (Oxford University Press 2014).
- Russell M, ‘Global Human Rights Sanctions: Mapping Magnitsky Laws — The US, Canadian, UK and EU Approach’ (European Parliamentary Research Service Briefing PE 698.791, November 2021).
- Ruys T and Ryngaert C, ‘Secondary Sanctions: A Weapon out of Control? The International Legality of, and European Responses to, US Secondary Sanctions’ (2019) 89 *British Yearbook of International Law* 12.
- Schmitt M N (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge University Press 2017).

Author’s note: This is a Phase II theory paper — a working preprint circulated for scholarly discussion. Its foundations lie in the author’s LL.M. dissertation, “Economic Warfare: The Unregulated War” (University of Worcester 2024), a separate and self-standing work that treats economic warfare, indirect proxy conduct and the Magnitsky instruments at length, and which is available in its own right. That dissertation is cited throughout the present paper as prior work rather than reproduced, and this paper is written to be read independently of it. The paper extends and, in its conclusion, revises the dissertation’s position: where the dissertation paired an eventual international charter with inspiring domestic legislation, this paper treats localised legislation as the primary remedy and the treaty as, at most, its downstream consequence. The “double veil” framing of Part 3 and the doctrinal mechanism of Part 5 (AI-conduct-triggered designation) are the author’s original contributions, offered as a model for legislative adaptation rather than as a statement of existing law. Comments are welcome.

— Christopher I. V. Farmer, LL.M. (University of Worcester) · Independent Researcher · civfarmer@gmail.com

Further reading and points of disagreement (July 2026 addendum)

Part 7 already concedes the strongest objections to designation as an instrument; this addendum anchors that concession to the literature a reviewer would expect and states why the thesis survives it. The empirical sanctions-scepticism canon — Hufbauer and colleagues’ comprehensive survey, Drezner’s “sanctions sometimes smart” reassessment, and the UN targeted-sanctions data assembled by Biersteker and colleagues, in which coercion succeeds in a clear minority of episodes — cuts against any claim that designation reliably changes a determined actor’s behaviour. This paper makes no such claim. The cascade argued in Part 4 operates through constraint and signal — raising the cost of access to indispensable market nodes and hardening a norm by attaching consequences — which is precisely the mode the targeted-sanctions data reports as most effective. The secondary-sanctions critique of Ruys and Ryngaert warns against jurisdictional overreach; the mechanism of Part 5 is drafted inside an established nexus for that reason. And the attribution difficulties that motivate Part 3 are the same ones the Tallinn Manual project documents for cyber operations, which is why conduct-triggered designation deliberately does not wait for state attribution. On the AI-governance side, the NIST AI Risk Management Framework is the natural domestic hook for the control-and-fault nexus developed in Part 5.

- Gary Clyde Hufbauer, Jeffrey J Schott, Kimberly Ann Elliott and Barbara Oegg, *Economic Sanctions Reconsidered* (3rd edn, Peterson Institute for International Economics 2007).
- Daniel W Drezner, ‘Sanctions Sometimes Smart: Targeted Sanctions in Theory and Practice’ (2011) 13 *International Studies Review* 96 <<https://doi.org/10.1111/j.1468-2486.2010.01001.x>>.
- Thomas J Biersteker, Sue E Eckert and Marcos Tourinho (eds), *Targeted Sanctions: The Impacts and Effectiveness of United Nations Action* (Cambridge University Press 2016) <<https://doi.org/10.1017/CBO9781316460290>>.
- Tom Ruys and Cedric Ryngaert, ‘Secondary Sanctions: A Weapon out of Control? The International Legality of, and European Responses to, US Secondary Sanctions’ (2019) 89 *British Yearbook of International Law* 12 <<https://doi.org/10.1093/bybil/braa007>>.
- Michael N Schmitt (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge University Press 2017) <<https://doi.org/10.1017/9781316822524>>.
- National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1, 26 January 2023; under revision as at 16 October 2026) <<https://www.nist.gov/itl/ai-risk-management-framework>> accessed 16 October 2026.