From Screen to Harm: Child Exploitation in the Digital Age

AI Disclosure Statement:

The author used an AI-assisted tool solely for formatting assistance and correction of citation formatting. All legal research, analysis, interpretation, arguments, and conclusions presented in this manuscript are the author's own work.

Abstract

Child exploitation was traditionally a bounded, physical problem, but the digital era has transformed it into something constant, algorithmically produced, and transnational. AI companion apps, recommendation systems, and chatbots now play a role in children's emotional lives, with documented cases of grooming, sexualised interaction, and psychological manipulation. This is evident in wrongful death litigation in American courts, where a federal court has, for the first time, accepted that a chatting emotional AI system can plausibly owe a duty of care to minors. These harms implicate rights guaranteed by the Convention on the Rights of the Child and clarified for the digital environment by General Comment No. 25, including protection from exploitation, privacy, and the best interests of the child.[1]

Yet international human rights law was built around states, not multinational companies that design, deploy, and profit from these technologies.[2] A child harmed in one country therefore has limited means to pursue a company based in another. This paper argues that the mismatch between where harm occurs, where corporate power sits, and where remedies can be found is a problem of global justice, not merely technology regulation. It leaves children in weaker states less protected despite holding identical rights on paper. Drawing on the UN Guiding Principles on Business and Human Rights, emerging guidance on child rights impact assessments, and comparative regulation, this paper asks which parts of this gap existing soft-law tools can close and which require new international instruments.

[1] Comm. on the Rts. of the Child, *General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment*, ¶¶ 4, 7, U.N. Doc. CRC/C/GC/25 (Mar. 2, 2021).
[2] U.N. Comm. on the Rts. of the Child, General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment, ¶¶ 12, 35, 112–13, U.N. Doc. CRC/C/GC/25 (Mar. 2, 2021).

Keywords: Child rights; Digital exploitation; General Comment No. 25; AI companion applications; Corporate accountability.

I. Introduction

In September 2025, the family of a thirteen-year-old girl from Colorado took the matter to federal court, filing a wrongful-death claim over a companion chatbot[3] that she had used for over a year, which had fostered an unhealthy emotional dependency and had not responded when she repeatedly expressed suicidal thoughts.[4] The chatbot itself was built by a company incorporated in Delaware, and it ran, in large part, on infrastructure controlled by a much bigger technology corporation. It reached a child in a state that had no law at the time directly regulating conversational AI.[5] Nobody had written the rules yet. What makes this case different from earlier ones is worth sitting with. It did not begin with a single unlawful image or message; the kind of thing child sexual abuse material statutes were written to catch. It began with a product working exactly as it was built to work. The harm that occurred was not a malfunction or a bad actor slipping through a gap but it was the ordinary, everyday operation of something designed, quite deliberately, to keep a child chatting with an AI chatbot. It is one of several such cases now moving through American courts,[6] and it captures, with unusual clarity, the problem this paper examines: a specific harm to a specific child, produced by a specific design choice, made by a specific company, in circumstances where the child's own government has limited practical reach over that company, and where the company's home government owes no clear, binding obligation under international law to prevent harm occurring abroad.

This paper is not an account of children and technology in general. It asks a narrower and more answerable question: when digital exploitation of a child is produced or worsened by the design of a transnational digital product, does human rights law give the child an effective route to accountability and remedy, and if it does not, what exactly is missing. There is now a considerable body of writing on children's digital rights, much of it prompted by the UN CRC's

[3] Complaint, Peralta v. Character Technologies, Inc., US District Court, District of Colorado (filed Sept. 2025).
[4] Garcia v. Character Technologies, Inc., 785 F. Supp. 3d 1157, 1160-63 (M.D. Fla. 2025).
[5] *Id.* at 1161-62.
[6] Shweta Watwe, *Character.AI, Google Agree to Settle Teen Chatbot Harm Lawsuits*, Bloomberg Law (Jan. 8, 2026), https://news.bloomberglaw.com/litigation/character-ai-google-agree-to-settle-teen-chatbot-harm-lawsuits [https://perma.cc/FN8P-RAQN] (last visited Sept. 21, 2026); Complaint, Peralta v. Character Technologies, Inc., US District Court, District of Colorado (filed Sept. 2025).

GC No. 25 of 2021.[7] Most of it addresses either the content of children's digital rights — privacy, participation, protection from harmful material — or the domestic regulatory response of individual states.[8] Less has been written about the structural question of where responsibility actually sits once the child, the corporation and the relevant infrastructure are spread across different countries, and about the shift, still under-theorized, from harms caused by illegal content that a state can order removed, to harms caused by a product's ordinary, lawful design.

The argument developed here is that the principal weakness in the existing framework is not a shortage of substantive rights for children — these are, on the page, unusually well developed — but the fragmentation of responsibility among three actors who are rarely found in the same country: the state where the child lives, the state where the controlling company is incorporated or headquartered, and the private company itself.[9] That fragmentation leaves a gap. The harm happens in one place. The technological power sits in another. And an effective remedy, if one exists at all, sits somewhere else again. The sections that follow test this argument against the evidence: the treaty text itself, the Committee's interpretive guidance, the business and human rights framework, comparative regulation, and the recent wave of AI companion litigation.[10] Along the way, this paper asks a harder question too. Does the evidence actually support the argument as I have framed it, or does it need to be refined?

The central research question is whether existing international human rights law provides an effective accountability and remedy framework where digital exploitation of children is generated or intensified by transnational digital platforms. Four further questions in this paper guide the discussion: how digital design, as distinct from digital content, produces new or

[7]UN Committee on the Rights of the Child, General Comment No. 25 (2021) on Children's Rights in relation to the Digital Environment, CRC/C/GC/25 (Mar. 2, 2021); Naomi Lott (ed.), *The Interdependence of the Convention on the Rights of the Child: "Understanding the Relationship of the Right to Play with Other Convention Rights"* (Springer 2025),; Emily Sylwander, *Mapping the Global Impact of General Comment No. 25*, London School of Economics and Political Science (2026), available at: https://researchonline.lse.ac.uk/id/eprint/137457/1/Mapping_the_global_impact_of_GC25_Sylwander_2026.pdf, [https://perma.cc/QTU4-3XTL] (last visited Aug. 14, 2026).

[8] Leanne Bowler, "The Digital Rights of Children and Youth: An Annual Review of Information Science and Technology (ARIST) Paper", *Journal of the Association for Information Science and Technology* (2026) (surveying privacy, data surveillance, age verification and participatory design as the dominant themes in children's digital rights scholarship).

[9] John Ruggie, *Protect, Respect and Remedy: A Framework for Business and Human Rights*, UN Doc. A/HRC/8/5, ¶ 3 (Apr. 7, 2008).

[10] Human Rights Council, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework*, U.N. Doc. A/HRC/17/31, princ. 25 & commentary (Mar. 21, 2011).

heightened forms of exploitation; which rights under the CRC and General Comment 25 are actually engaged, and what these instruments require of states; what jurisdictional doctrine governs a child's ability to reach a foreign company at all; and whether the UN Guiding Principles on Business and Human Rights, and the child-specific due diligence tools now being built on them,[11] offer a realistic basis for closing the accountability gap, or whether closing it needs new binding instruments the international system does not yet have.

The method is doctrinal. It works from the text of the Convention and its Optional Protocols, the Committee's General Comments, the Guiding Principles, and comparative domestic and regional regulation, and it uses a small number of carefully chosen examples to test rather than merely illustrate the analysis. Empirical claims — about litigation outcomes, or the concentration of major platforms in particular countries — are kept separate from the normative legal claims about what the law actually requires, since the two are easily, and wrongly, run together in this area.

II. The Shift from Harmful Content to Harmful Design

Child exploitation has no single settled definition across international instruments. The Convention, together with its Optional Protocol addressing the sale of children, child prostitution, and child pornography, covers sale, prostitution, pornography, and closely related forms of sexual exploitation.[12] Article 32 addresses economic exploitation, and Article 36 supplies a residual protection against "all other forms of exploitation prejudicial to any aspects of the child's welfare."[13] That residual clause matters here: it was drafted broadly enough to cover harms the drafters could not have foreseen, and it is the textual route by which newer, technologically produced harms fall within the Convention's protection without requiring a new treaty.

[11] Jelena Bäumler et al. (eds.), European Yearbook of International Economic Law 2024 (Springer 2025).
[12] Republic of The Gambia, *Combined Fourth to Seventh Reports Submitted by The Gambia Under Article 44 of the Convention, Due in 2021*, U.N. Doc. CRC/C/GMB/4-7 (May 20, 2024), https://www.ecoi.net/en/file/local/2110656/G2408345.pdf [https://perma.cc/5LCH-ZUGC] (last visited Aug. 14, 2026).
[13] Convention on the Rights of the Child art. 36, Nov. 20, 1989, 1577 U.N.T.S. 3.

General Comment No. 25 gives that residual protection real content for the present day.[14] The Committee treats the digital environment as an extension of a child's ordinary life rather than a separate space with its own rules, and insists that a child's rights must be respected, protected and fulfilled online exactly as they are offline.[15] That matters doctrinally. This shuts down a convenient excuse: the idea that a state's responsibility ends right at the edge of a platform.

For the present purposes, it helps to separate two kinds of harm that the literature tends to combine together. The first is content based exploitation in which there is an unlawful image, message or video, harmful no matter how the child comes across it.[16] The legal answer here is straightforward enough: take it down, and where the source can be traced, prosecute. This is the model behind the Optional Protocol and the Council of Europe's Lanzarote Convention on protection of children against sexual exploitation and abuse.[17] It is also the logic behind the cybercrime cooperation built into the Budapest Convention.[18] All three instruments assume the same thing: a discrete unlawful item that has a discoverable point of origin, something you can point to and trace back.

Design based exploitation is a different problem altogether. The harm here does not live in a piece of content. It lives in the ordinary, intended working of a system. A recommender algorithm designed simply to keep a child scrolling. A chatbot persona built to make a child feel it is emotionally real — as real as a human. A notification designed to interrupt whatever the child happens to be doing. Or simply the absence of any real age check at all — so a product that's built for adult interaction ends up chatting with a nine-year-old, sometimes a thirteen-year-old. None of this is content-based harm, nor can the authorities remove it, because there is nothing to delete.[19]

---

[14] Comm. on the Rts. of the Child, *General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment*, ¶¶ 4, 7, U.N. Doc. CRC/C/GC/25 (Mar. 2, 2021); Christian Djeffal, *Children's Rights by Design and Internet Governance: Revisiting General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment*, 11 Laws 84 (2022).

[15] *Id,*

[16] United Nations Children's Fund, Protecting Children from Violence and Exploitation in Relation to the Digital Environment: Policy Brief *2* (2024).

[17] Council of Europe Convention on the Protection of Children Against Sexual Exploitation and Sexual Abuse, CETS No. 201, art. 3 (2007).

[18] Council of Europe, Convention on Cybercrime, CETS No. 185, 2001 (entered into force July 1, 2004).

[19] European Commission, *Guidelines on Measures to Ensure a High Level of Privacy, Safety and Security for Minors Online, Pursuant to Article 28(4) of Regulation (EU) 2022/2065*, C/2025/5519 (Oct. 10, 2025).

The Character Technologies litigation makes the point much clearer than any fictional case could. The claim was never that the platform hosted unlawful material. It was the product design itself, an interference engineered to sustain emotionally realistic engagement, created a foreseeable risk to a minor using it. And a federal court in Florida found that theory, product liability and negligence together, plausible enough to survive a motion to dismiss.[20]

Regulators have begun to catch up with this shift. The European Commission's 2025 guidelines on Art. 28(1) of the Digital Service Act, which requires platform accessible to minor to adopt[21] "*appropriate and proportionate measures*" to protect their "*privacy*", "*safety*" and "*security*",[22] reads that obligation into design requirements, restrictions on profiling-based recommender systems for minors, default-private settings, restrictions on interface features engineered to manipulate user behavior, including infinite scroll, and a specific rule that AI chatbots must not be switched on by default for children or presented in ways that risk being mistaken for a human.[23] This is, in substance, a regulatory acceptance that relevant unit of control is no longer the individual item of content but the system's design, and GC. No. 25 had already anticipated it, calling on states to require "child rights by design" from any provider whose products children are likely to use.[24]

The point matters for the argument of this paper in two ways. Design-based harms are harder to trace to a single point of origin than content-based harms, because a recommender system's output is the product of code, training data and commercial incentives that may be developed and refined across several countries at once, none of which is necessarily the country where the child actually experiences the harm.[25] And the established child-protection instruments — the Optional Protocol, Lanzarote, Budapest — were built for the older, content-based paradigm and do not squarely reach design. General Comment No. 25 and the DSA guidelines begin to close that gap,

[20] *Supra* note 4, at 1175-80 (design-defect and negligence claims permitted to proceed).
[21] European Commission, *Guidelines on Measures to Ensure a High Level of Privacy, Safety and Security for Minors Online, Pursuant to Article 28(4) of Regulation (EU) 2022/2065* (Oct. 10, 2025) (on file with author) (last visited Sept. 21, 2026).
[22] *Id.*
[23] *Id*.
[24] *Supra* note 14.
[25] Anna-Katharina Meßmer & Martin Degeling, *Auditing Recommender Systems—Putting the DSA into Practice with a Risk-Scenario-Based Approach*, arXiv:2302.04556 (2023), [https://perma.cc/8E9V-UDRN] (last visited Sept. 21, 2026).

but one is interpretive guidance binding no one directly, and the other is regional, reaching platforms outside the EU only to the extent those platforms choose to serve EU users.[26]

III. Child's Rights in the Digital Environment

The starting point remains the CRC, ratified by 196 states and signed but not ratified by the US — a distinction examined more closely below.[27] Several provisions bear directly on digital exploitation. Art. 3 requires that "the best interests of the child shall be a primary consideration in all actions concerning children" — language the Committee reads as extending to the design choices of businesses whose products children use. Art. 6 protects the "right to life, survival and development," engaged wherever platform design contributes to serious psychological harm or, in the worst documented cases, to suicide.[28] Art. 12 protects the "child's right to be heard," which General Comment No. 25 extends to require that children's own experience of digital products should shape how those products are designed and regulated, not only adult assumptions about what is safe for them.[29] Art. 16 protects against arbitrary interference with privacy, directly engaged by profiling and behavioral targeting.[30] Art.19 requires states "to take all appropriate measures to protect children from violence, abuse or exploitation, language the Committee reads expansively enough to cover harm mediated by a digital service a child uses without direct adult supervision."[31] Articles 34, 35 and 36 address "*sexual exploitation, trafficking and other exploitation*" respectively, with Article 36 again supplying the residual category that catches novel digital harms.[32]

Two further instruments matter. The Optional Protocol on the "sale of children, child prostitution and child pornography" requires states to criminalize these practices and to cooperate in prevention, detection and prosecution, and its enforcement architecture — built around traditional inter-state criminal cooperation — has since been read to extend to online-facilitated

[26] European Commission, *Guidelines on the Protection of Minors under Article 28(4) of Regulation (EU) 2022/2065* (July 14, 2025) (on file with author) (last visited Sept. 21, 2026).
[27] U.S. Cong. Rsch. Serv., *The United Nations Convention on the Rights of the Child*, R40484 (2015).
[28] *Supra* note 4, at 1163 (allegations of prolonged engagement and suicidal ideation).
[29] *Supra* note 14.
[30] Convention on the Rights of the Child art. 16, Nov. 20, 1989, 1577 U.N.T.S. 3.
[31] *Supra* note 14.
[32] Convention on the Rights of the Child arts. 34–36, Nov. 20, 1989, 1577 U.N.T.S. 3; John Tobin, *Article 36: Protection Against All Other Forms of Exploitation*, in *The U.N. Convention on the Rights of the Child: A Commentary* 1402, 1402–19 (John Tobin ed., 2019).

sale and sexual exploitation, though it was not written with corporate regulatory obligations in mind.[33] In force since 2014, the Optional Protocol on a Communications Procedure lets a child — or someone acting on their behalf — bring a complaint to the Committee claiming a state has breached their Convention rights. Two conditions apply: the respondent state must have separately ratified the Protocol, and the child must first have exhausted all available domestic remedies[34] — a real limitation discussed further below.

General Comment. 16 of 2013, on the obligation of state regarding the business sector impact on the children's rights, is the Committee's clearest statement on the corporate dimension of this problem.[35] It confirms that nonetheless the Convention "does not create direct obligations for business enterprises," but holds that states parties must "respect, protect and fulfil children's rights in relation to business activity," including by regulating businesses domiciled within their own jurisdiction and by taking measures, where they are able, to prevent extraterritorial abuse by companies over which they have some degree of control or influence — through incorporation, headquarters or listing status.[36] The formulation is cautious. The Committee stops short of asserting a freestanding extraterritorial duty independent of any jurisdictional link, but this is nonetheless the clearest statement within the framework that a state's obligations to children do not end at its own border the moment a company under its effective control causes harm abroad.[37]

General Comment No. 25 builds directly on this for the digital environment, calling states to require digital providers to carry out child-right due diligence, to present information to children

[33] Optional Protocol to the Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography arts. 2–5, May 25, 2000, 2171 U.N.T.S. 227 [hereinafter OPSC]; Comm. on the Rts. of the Child, *Guidelines Regarding the Implementation of the Optional Protocol to the Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography*, U.N. Doc. CRC/C/156, ¶¶ 9–12, 16–17 (2019); Comm. on the Rts. of the Child, *Concluding Observations on Georgia's Initial Report Under the Optional Protocol to the Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography*, U.N. Doc. CRC/C/OPSC/GEO/CO/1, ¶¶ 29–30 (2019).

[34] Optional Protocol to the Convention on the Rights of the Child on a Communications Procedure, arts. 5, 7, Dec. 19, 2011, 2983 U.N.T.S. 135; UN Treaty Collection, *OPIC*, Status of Treaties (showing entry into force Apr. 14, 2014).

[35] GC No. 16: *CRC, General Comment No. 16 (2013), U.N. Doc. CRC/C/GC/16 (Apr. 17, 2013).*

[36] UN CRC, GC No. 16 (2013) on "State Obligations regarding the Impact of the Business Sector on Children's Rights," CRC/C/GC/16 (Apr. 17, 2013).

[37] Federico Casolari, *GC No. 25 on "Children's Rights in Relation to the Digital Environment: Implications for Children's Right to Privacy and Data Protection in Africa"*, 24 **Hum. Rts. L. Rev.** 1, 8–10 (2024); CRC, *GC No. 16 (2013) on "State Obligations Regarding the Impact of the Business Sector on Children's Rights*," U.N. Doc. CRC/C/GC/16, ¶¶ 43, 66–67 (2013).

in a form they can actually understand, to "*prohibit profiling of children for commercial purposes*" absent a demonstrated best-interests justification, and to ensure children have access to remedy mechanisms that are genuinely usable by a child.[38] It remains interpretive guidance rather than an independent source of binding obligation, but General Comments carry real interpretive weight in the Committee's own review of state reports, and states are expected to account for their implementation.[39]

IV. Where Responsibility Breaks Down

The argument can now be stated more precisely. A child in one country uses a digital product controlled by a company incorporated in a different country. The engineering, the training, the commercial decisions behind that product might be made in another country entirely, for example, Apple Inc. is incorporated in the United States while much of its hardware is manufactured in China, so the companies are spread across a whole set of jurisdictions through cloud infrastructure nobody outside the company can fully see. Three persons are potentially responsible here. And none of the existing instruments divides that responsibility between them with any real precision.

The child's own state carries the primary obligation under the Convention. It is supposed to protect the child within its border. But its actual tools, enforcement action, blocking orders, going after intermediaries, only work if the company has something to lose there. Some presence. Some assets. Some commercial reason to bother complying at all.[40] Where a platform has no local subsidiary, no local revenue that can be reached, and no interest in maintaining local market access, a state's formal jurisdiction over its own territory does very little to give it actual control over the source of the harm.[41] The European Union avoided exactly this problem within its own market by giving the Digital Services Act extraterritorial reach — the Act applies to any provider

[38] *Supra* note 14.
[39] CRC, *GC No. 25 (2021) on "Children's Rights in Relation to the Digital Environment*," U.N. Doc. CRC/C/GC/25, ¶¶ 4, 7 (2021), Federico Casolari, General Comment No. 25 on Children's Rights in Relation to the Digital Environment: Implications for Children's Right to Privacy and Data Protection in Africa," 24 *Human Rights Law Review* 1, 4–5 (2024); CRC arts. 44–45, Nov. 20, 1989, 1577 U.N.T.S. 3.
[40] John Ruggie, Protect, Respect and Remedy: A Framework for Business and Human Rights, U.N. Doc. A/HRC/8/5, ¶¶ 18–21 (2008).
[41] *Id*.

serving EU users regardless of where that provider is established[42] — but that solution is available only to a bloc large enough to make compliance commercially worthwhile for a foreign platform, and it cannot easily be copied by a state without comparable market size.

The corporate home state owes the foreign child no direct obligation under the Convention at all, since the Convention's obligations run to "*each child within [the state's] jurisdiction*," language generally read, in the Committee's practice and in comparable human rights jurisprudence, as territorial or as requiring effective control, not as a general duty to prevent harm anywhere in the world caused by a company that happens to be domiciled there. General Comment No. 16's language that a state "*should*" act where it has a reasonable link to the conduct is an expectation the Committee has articulated, not a settled rule states have accepted as binding.[43]

The company itself is not a direct duty bearer under the Convention. Not at all. The GP(Guiding Principles) do impose something called a "responsibility to respect human rights," and this exists "independently of whether any state enforces it."[44] But it is a standard of expected conduct, grounded in social expectation. It is not a legal liability a child can actually invoke before an international body. There is a real difference between the two, and it matters here.

Where liability does arise, it arises through domestic law. Negligence. Product liability. Consumer protection. Data protection. Whatever a court can get its hands on, really, applied by whichever court can establish jurisdiction, and whose judgment can, in practice, actually be enforced against someone.

The Character Technologies litigation shows what becomes possible when all of this happens to line up. A child user sat within the jurisdiction of a court that could reach a Delaware incorporated company, one with substantial assets in the United States.[45] That is not a small thing. But where the child is not a resident of a country with anything like comparable product

[42] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), art. 2(1), recital 7, 2022 O.J. (L 277) 1, 18.
[43] Paula Gerber, Joanna Kyriakakis & Katie O'Byrne, General Comment 16 on State Obligations Regarding the Impact of the Business Sector on Children's Rights: What Is Its Standing, Meaning and Effect?, 14(1) *Melbourne Journal of International Law* 93, 93–128 (2013).
[44] BSR, "Child Rights Impact Assessments in Relation to the Digital Environment" 4 (UNICEF/BSR, Jan. 2025).
[45] *Supra* note 4, at 1165-67 (personal jurisdiction over Delaware-incorporated defendant).

liability doctrine, and the platform has no presence there either, no equivalent route exists right now. None.

The result is not an absence of rights, as Article 36 and General Comment No. 25 plainly cover the harm described above but there is an absence of any institution which can turn that right into a remedy against the person who actually controls the relevant design choice, in the place where the child actually is.[46] That is the precise sense in which responsibility is divided: it covers on paper very broadly, but the machinery for enforcing it against transnational private power is narrow and unevenly available.

It is worth separating several ideas that are often run together in this area. State responsibility under general international law arises from conduct attributable to a state, not from purely private conduct merely because the private actor happens to be incorporated there.[47] State due diligence is the "positive obligation" to take reasonable regulatory steps against foreseeable harm within the state's jurisdiction or effective control — the obligation the Committee invokes in GC 16 and 25. Direct corporate liability under international law remains largely undeveloped: no binding treaty currently imposes "human rights obligations on corporations enforceable" before an international body. Parent-company and platform liability are domestic-law doctrines that vary from one legal system to the next and determine whether a child can reach a foreign parent through a local subsidiary at all. Intermediary liability, concerning whether a platform is responsible for third-party content it hosts, is increasingly beside the point for design-based harm, since the platform is not hosting someone else's content but generating the interaction itself, as in the chatbot cases.[48]

On the specific question of whether a state must prevent extraterritorial harm caused by a corporation domiciled within it, the position under general international law is unsettled rather than resolved either way. Regional human rights jurisprudence recognizes that a "state's obligations can extend beyond its territory where it exercises effective control over territory or a

[46] *Supra* note 4.
[47] International Law Commission, Draft Articles on Responsibility of States for Internationally Wrongful Acts, with Commentaries, arts. 2, 8, U.N. Doc. A/56/10, at 34, 47–48 (2001); Military and Paramilitary Activities in and Against Nicaragua (*Nicar. v. U.S.*), Judgment, 1986 I.C.J. Rep. 14, ¶¶ 109–115 (June 27); James Crawford, State Responsibility: The General Part 113–24 (Cambridge University Press 2013); Application of the Convention on the Prevention and Punishment of the Crime of Genocide *(Bosn. & Herz. v. Serb. & Montenegro)*, Judgment, 2007 I.C.J. Rep. 43, ¶¶ 396–406 (Feb. 26).
[48] *Id*.

person abroad," but this doctrine developed in the context of military occupation and detention, and no binding tribunal has extended it into a general duty to regulate the extraterritorial conduct of domestically incorporated companies toward foreign children.[49] This qualifies the paper's central claim in an important way. The accountability gap is not simply a failure to enforce some rule that already exists and is just being ignored. Part of the gap is more basic than that. The underlying obligation, the idea that a corporate home state owes some duty of due diligence over harm its companies cause abroad, is itself still emerging soft law. It has not settled into binding law yet. So any real reform proposal has to deal with that prior question too, not just the question of enforcement.

This has particular force for the United States, home to most of the platform and AI companion companies at issue here. It is worth being precise about the country's actual legal position, since the claim that the US has "never signed" the Convention is a common but inaccurate shorthand. The US signed the Convention on 16 February 1995, through then-Ambassador Madeleine Albright, but the Clinton administration never TRANSMITTED it to the Senate for "advice and consent," and no administration has done so since; the US is therefore signatory but not a party, and remains the only "UN member state that has not ratified it,"[50] a position unchanged since Somalia's ratification in 2015.[51] Signature without ratification carries a narrower consequence under Art 18 of the VCLT: "a signatory must refrain from acts that would defeat the treaty's object and purpose pending ratification, but is not bound by its substantive terms." The US has, however, ratified both OP(optional prototcol) to the Convention, including the one on sale, prostitution and pornography, following Senate advice and consent in 2002,[52] so it carries binding obligations in that narrower area even though the parent Convention does not bind it. This matters directly for the case studies that follow: the state most likely to be the corporate home of an AI companion company is not itself bound by the Convention's general framework,

[49] U.N. Human Rights Council, Guiding Principles on Business and Human Rights: Implementing the United Nations Protect, Respect and Remedy Framework, U.N. Doc. A/HRC/17/31, princs. 1–2 (2011; Committee on Economic, Social and Cultural Rights, *General Comment No. 24 (2017) on State Obligations Under the International Covenant on Economic, Social and Cultural Rights in the Context of Business Activities*, U.N. Doc. E/C.12/GC/24, ¶¶ 25–27 (2017),; *Banković and Others v. Belgium and Others*, App. No. 52207/99, 2001-XII Eur. Ct. H.R., ¶¶ 59–61 (2001).

[50] Stacey Barell Steinberg, "Beyond Sharenting," 99 *Southern California Law Review* (2026), available at: https://southerncalifornialawreview.com/2026/04/22/beyond-sharenting-2/, https://perma.cc/B2QM-W9ZB (last visited Aug. 14, 2026).

[51] *Supr*a note 27.

[52] *Supra* note 27, at 3.

including Article 36 or General Comment No. 25, though it remains bound by the Optional Protocol's narrower sexual-exploitation obligations and, separately, by its own extensive domestic tort, consumer-protection and product-liability law — which is, in practice, the actual source of the remedies now being litigated against Character Technologies.

V. Remedies in Practice

Whether a child in this position has a realistic route to remedy, rather than a merely formal one, depends on which channel is open. A domestic court in the child's own state needs personal jurisdiction over the corporate defendant — something that's often unavailable without a local subsidiary in that country. And even if the jurisdiction is established it is really difficult to enforce the judgment against the foreign entity. A domestic court in the corporate home state can offer a genuine remedy, as the Character Technologies litigation shows, but usually only to the persons who can establish standing and jurisdiction there in the first place — easiest for domestic plaintiffs, considerably harder for a foreign child with no resources for cross-border litigation. A regulatory complaint to a data-protection or child-protection authority can compel a change in behavior, as the "Digital Services Act's" Art. 28 enforcement shows, but rarely delivers individual compensation and only binds platforms within the regulator's own reach. A communication to the Committee under the "Optional Protocol on a Communications Procedure" allows a child to complain about a state's failure, not directly about a company's conduct, and only against states that have ratified that Protocol, after domestic remedies are exhausted — "*a slow process producing recommendations rather than enforceable orders,*" with no way to bind a company domiciled in a third state that never appears before the Committee at all. Corporate grievance mechanisms, encouraged under the Guiding Principles and adopted voluntarily by some large platforms, can move faster, but are not independent, are not bound by Convention standards, and their outcomes cannot be judicially enforced.

None of these channels, on its own, gives a child a real route to remedy against a foreign-domiciled platform's design choices, unless the child happens to live in, or can reach the courts of, the platform's home country. The Character Technologies settlements of January 2026, and the separate suits brought by the Kentucky and Pennsylvania authorities, show what domestic regulatory and law-enforcement tools can achieve where a state has both territorial

jurisdiction and the institutional capacity to litigate against a well-resourced technology company.[53] That capacity is not evenly spread across the 196 states "party to the Convention," and this unevenness is the paper's underlying global justice claim: the same design flaw, producing the same category of harm, is remediable for a child in one country and effectively irremediable for a child in another — not because their underlying right differs, but because of where they and the responsible company happen to be, and what legal machinery their own state can bring to bear.

Four regulatory models illustrate different ways of closing part of this gap; none built as an international human rights instrument but each bearing directly on the accountability question. The European Union's approach, through the Digital Service Act and its 2025 guidelines, is the most developed one so far as it applies to any platform serving EU users regardless of where the company is incorporated, imposes design-specific obligations, and it is backed by real enforcement powers, including fines as well. But it only protects children in or accessing services from the EU, and does nothing for a child elsewhere using that same platform. The United Kingdom's "Age-Appropriate Design Code," made under the "Data Protection Act 2018," was an earlier and fairly influential model, requiring services likely to be accessed by "children to apply high privacy settings by default."[54] It has shaped similar codes elsewhere, but its binding force also stops at the edge of UK regulatory reach. Regulation in the United States, by contrast, has developed far more unevenly — largely through state legislatures rather than any comprehensive federal statute. California's SB 243, effective from 1 January 2026, was the first state law to impose obligations specific to companion chatbots — AI disclosure, crisis-referral protocols, mandatory break reminders for minors, and a private right of action — and New York's comparable law took effect shortly before it, on 5 November 2025.[55] Australia has taken a fourth and structurally different route: rather than regulating how a platform is designed for children, "*the Online Safety Amendment (Social Media Minimum Age) Act 2024*" excludes children under sixteen from covered platforms altogether, placing the burden on the platform to

[53] Office of the Attorney General, Commonwealth of Kentucky, *AG Coleman Sues AI Chatbot Company for Preying on Children* (Jan. 8, 2026) (on file with author) (last visited Sept. 21, 2026); Commonwealth of Pennsylvania, Governor's Press Office, *Shapiro Administration Sues Character.AI Over Fake Medical Claims* (May 5, 2026) (on file with author) (last visited Sept. 21, 2026).

[54] Data Protection Act 2018, c. 12, § 123 (UK); Information Comm'r's Off., *Age Appropriate Design: A Code of Practice for Online Services*, at 61 (2020).

[55] California Senate Bill 243 (Companion Chatbots), 2025 Cal. Stat., effective Jan. 1, 2026; New York AI Companion Models Law, 2025 N.Y. Laws, effective (Nov. 5, 2025).

take reasonable, auditable steps to verify age and remove existing accounts held by under-16 users once the obligation took effect on 10 December 2025.[56] This is a categorically different strategy from design regulation: it does not ask what a safe product for children would look like, but removes the child from the product's reach entirely, on the premise that no amount of design mitigation makes an engagement-optimized platform genuinely safe for that age group. It shares the same jurisdictional limit as the other three models, however — the obligation binds platforms only in respect of Australian account-holders, so a child outside Australia interacting with the same company gains nothing from it, and the safety Commissioner's own guidance leaves considerable discretion to platforms over what counts as a reasonable verification method, which is itself likely to generate further design-based disputes of exactly the kind this paper discusses.

None of these four models, on their own or taken together, adds up to an international framework that binds a company regardless of where the child has a place to be. Each one is a domestic or regional response. Each of the Four model works because that particular jurisdiction had the market size, or the regulatory tradition, or the litigation culture and corporate presence, to make it bind.

This is where the paper's main global justice argument really gets interesting. Effective protection right now tracks regulatory and economic capacity. It does not track how severe the risk to the child is, nor the strength of the underlying right, since that right is, on paper, formally identical for every child, wherever they happen to live.

The Character Technologies litigation deserves a closer look. It is well documented through primary sources, and it tests the argument directly rather than just illustrating it from a distance.

The claim concerned a fourteen-year-old user in Florida. According to the complaint, he developed an intense emotional attachment to a chatbot persona over several months, and in February 2024 he died by suicide. His mother's suit alleged product liability, negligence and wrongful death. She sued the platform's developer, and alongside it, as a co-defendant, a much

[56] Online Safety Amendment (Social Media Minimum Age) Act 2024 (Cth) (Act No. 127 of 2024), effective Dec. 10, 2025; eSafety Commissioner (Australia), "Statement on the Online Safety Amendment (Social Media Minimum Age) Act 2024" (2025).

larger technology corporation with a substantial commercial and technical relationship to the company.[57]

In May 2025, the presiding federal judge rejected the argument that chatbot outputs were protected speech immune from suit, allowing the product-liability and negligence claims to proceed — a ruling widely treated as the first to establish that a conversational AI system can plausibly "owe a duty of care to a minor."[58] The case, together with related suits from Colorado, New York and Texas, settled on confidential terms in January 2026.[59] Two things follow. A remedy was genuinely available, but only because claimant and defendant were both within a single, well-resourced legal system with product-liability doctrine capable of being adapted to a new technology. And the theory of liability the court accepted was explicitly about design — the persona's engineered emotional realism, not any discrete unlawful content.[60]

The subsequent state actions extend the picture. "Kentucky's Attorney General" brought the first state suit against Character Technologies in January 2026, alleging that the platform "*preyed on children and led them into self-harm*," relying on consumer-protection and data-privacy law rather than tort,[61] and Pennsylvania's Department of State separately sued the company in May 2026, alleging its chatbots impersonated licensed medical professionals.[62] These show state authorities using consumer-protection statutes as an alternative accountability route, but they remain confined to American sub-national jurisdiction and offer no template usable by a child in a country without an equivalent office, budget or statutory framework covering AI systems.

The EC's July 2025 guidelines on Art. 28(1), by contrast, are a design-based, preventive intervention of exactly the kind General Comment No. 25 recommends, and they were possible only because the EU's market size gave it leverage over platforms with no EU corporate presence at all.[63] This suggests that the accountability gap identified above is not an unavoidable feature of transnational digital harm. It is closed wherever a jurisdiction has enough regulatory leverage,

[57] *Supra* note 4, at 1160-63.
[58] *Supra* note 4, at 1168-75.
[59] Watwe, *Supra* note 6.
[60] *Supra* note 4.
[61] Commonwealth of Kentucky v. Character Technologies, Inc., Complaint, Franklin Cir. Ct., Ky. (filed Jan. 8, 2026); Office of the Attorney General, Commonwealth of Kentucky, supra note 53.
[62] Commonwealth of Pennsylvania, Dep't of State, State Bd. of Medicine v. Character Technologies, Inc., No. 220 M.D. 2026 (Pa. Commw. Ct., filed May 1, 2026); Governor's Press Office, supra note 53.
[63] *Supra* note 4.

and remains open wherever it does not — which is the premise for the reform discussion that follows.

VI. What Reform Is Actually Possible

Some parts of this gap can be narrowed through existing tools, without new treaty text. States could make child rights impact assessments mandatory for high-risk digital products likely to be used by children, building on guidance already developed jointly by the UN human rights office and UNICEF for applying the Guiding Principles to the digital environment,[64] reinforced through the Committee's own reporting process. The Digital Services Act's design-specific obligations offer a template other large markets could adopt unilaterally, since the Act itself is a regional instrument applied extraterritorially rather than a product of new international law.[65] Wider ratification of the Optional Protocol on a Communications Procedure, currently in force for fewer states than the Convention itself, would at minimum give more children the formal option of complaining about their own state's failure to regulate, even though this does not reach the company directly. And more states could adopt a private right of action modelled on California's statute, giving children and families direct standing against AI companion developers without depending on prosecutorial discretion. Australia's age-exclusion model offers a further, distinct template — one that trades the complexity of ongoing design regulation for a bright-line eligibility rule, at the cost of depending entirely on the reliability of the age-verification technology used to enforce it.[66]

A more ambitious existing tool, and one still working through its own teething problems, is "*the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*," opened for signature in September 2024 as the first legally binding international treaty addressing AI systems as such, whose Art. 18 specifically requires "*parties to take due account of the needs and vulnerabilities of children in implementing the*

[64] *Supra* note 21.
[65] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), arts. 2(1), 34(1), 35(1)(a), (d), recitals 79–81, 2022 O.J. (L 277) 1.
[66] Cal. Bus. & Prof. Code §§ 22601–22606 (West 2025); Australian Gov't, *Age Assurance Technology Trial: Final Report* (2025); eSafety Comm'r, *Social Media Minimum Age Regulatory Guidance* (2025); Comm. on the Rts. of the Child, *General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment*, U.N. Doc. CRC/C/GC/25, ¶¶ 37–40, 53–54 (2021).

Convention."[67] It illustrates, more clearly than anything else discussed in this paper, how slowly a genuinely binding instrument moves even once the text itself exists: as of mid-2026 the European Union is the only party to have deposited an instrument of ratification, and Convention will not enter into force until five parties, at least three of them Council of Europe member states, have done so.[68] It is also narrower than it first appears for present purposes. Article 3(3) exempts research and development activity not yet available for use, which is precisely the stage at which a chatbot persona's engagement-optimized design is fixed, and the Convention regulates AI systems generally rather than creating the extraterritorial corporate-home-state duty toward foreign children that this paper identifies as the specific missing piece.[69] The "UN Global Digital Compact," adopted by consensus alongside the "Pact for the Future in September 2024", points in the same direction at a lower level of legal force: it commits states to strengthen legal and policy frameworks protecting children in the digital space and expressly invokes the Convention and GC No. 25 as the governing standard.[70] But it is a "political declaration of the General Assembly," not a treaty, and it creates nothing a child could invoke before any body. Taken together, the Framework Convention and the Global Digital Compact show real and recent movement in the direction this paper argues is necessary, but neither yet closes the specific gap identified above, and the Framework Convention's slow ratification record is itself evidence of how much political effort separates a binding text from a binding obligation actually in force.

Other parts of the gap genuinely require new international law, because nothing currently reaches this far. A treaty-based extraterritorial due-diligence obligation on the corporate home state — turning General Comment No. 16's "*should*" into a binding "*must*" — would need either a new "optional protocol to the Convention" or a freestanding instrument on business and children's digital rights, since a General Comment cannot itself create binding obligations beyond the treaty it interprets. A requirement that providers of AI companion products aimed at minors designate a locally reachable legal representative for cross-border complaints has some precedent in the Digital Services Act's requirement that non-EU providers appoint an EU representative, but has not been generalized into an international norm. An international complaint mechanism able to

[67] Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225, opened for signature Sept. 5, 2024, art. 18.
[68] *Supra* note 67, art. 30.
[69] *Supra* note 67, art. 3(3).
[70] UN General Assembly, Pact for the Future, Global Digital Compact and Declaration on Future Generations, GA Res. 79/1 (Sept. 22, 2024), annex I (Global Digital Compact).

receive communications about corporate conduct directly, rather than only about state failure, would be a genuine institutional innovation, since the Committee's mandate is to review state, not corporate, compliance. And without arrangements for enforcing judgments obtained by children against foreign corporate defendants, even a remedy secured in the corporate home state may remain difficult to actually access.

None of this should be presented as easily achievable. A treaty-based extraterritorial obligation faces the same political resistance that has stalled a broader binding "business-and-human-rights treaty": the UN working group mandated to draft one has been negotiating since 2014 and had completed eleven sessions without agreement on a final text by the time of its most recent meeting in October 2025, with a twelfth session scheduled for October 2026,[71] sharpened by the fact that the United States, the state most relevant to AI companion regulation, is not even a party to the underlying Convention.[72] A corporate-facing complaint mechanism would be a significant departure from the state-centered structure of the "UN human rights treaty system", and nothing in current Committee practice suggests the political will for it exists yet. Mandatory impact assessments, achievable without new treaty law, risk becoming a paper exercise unless paired with independent audit and a real sanction for failure. And a requirement to designate a local representative is only as effective as the company's own commercial incentive to comply, which does little for children in markets too small to matter to it — reproducing, rather than resolving, the underlying imbalance of leverage.

These difficulties don't defeat the argument; if anything, they sharpen it. This paper isn't claiming that international law can simply close the accountability gap once someone finds the political will. What it offers instead is a more careful position. Some parts of the gap can be closed now, through soft law that's already there and through comparative regulation spreading in countries from one to the next. Other parts are genuinely stuck, and they will stay stuck until new binding instruments exist — something current state practice gives little reason to expect any time soon.

[71] UN Human Rights Council, Res. 26/9 (June 26, 2014); OHCHR, Eleventh Session of the Open-Ended Intergovernmental Working Group on Transnational Corporations and Other Business Enterprises with respect to Human Rights (20-24 Oct. 2025).

[72] Convention on the Rights of the Child, Nov. 20, 1989, 1577 U.N.T.S. 3; United Nations Treaty Collection, Status of Treaties, Convention on the Rights of the Child (showing United States as signatory but not party).

That's a qualified conclusion, not an abandoned one. The gap is genuinely real. It's best understood as responsibility scattered across those who are duty bearers and who are rarely in the same country, let alone the same room. But closing it fully isn't just a matter of political will positioned at one missing rule. Part of the underlying obligation — the idea that a corporate home state must exercise real due diligence over harm abroad — is itself still soft law, still forming. Turning it into something harder means facing the same old arguments about sovereignty and fairness between rich and poorer states that have slowed down business and human rights treaty making for years.

VII. Conclusion

Digital exploitation of children is now well documented, covered by Art. 36 of the Convention and General Comment No. 25, and increasingly the subject of litigation. What has received far less attention is the structural reason a formally comprehensive set of children's rights which translates so unevenly into actual protection: responsibility for design-based digital harm is fragmented across the child's own state, the corporate home state and the company itself, and no existing instrument assigns or enforces that responsibility with any precision across borders. The recent AI companion litigation, the EU's Art. 28 guidelines and Australia's age-exclusion law show that the gap can be closed where a jurisdiction has enough regulatory or judicial leverage over the relevant company; the same litigation and regulation also show, by their very territorial limits, that children outside such jurisdictions remain comparatively unprotected despite holding identical rights. That is why this is properly a problem of global justice rather than simply of technological policy: unequal access to remedy tracks unequal state capacity and unequal proximity to concentrated corporate power, not unequal entitlement under international law. A workable reform agenda has to proceed on two fronts at once — extending what is already achievable through child rights impact assessments, wider ratification of the Communications Procedure, and the diffusion of comparative regulation, while being honest that a binding extraterritorial due-diligence obligation and a corporate-facing international remedy mechanism remain, for now, aspirations that current state practice does not yet support, and whose achievement will depend on the same contested politics that have limited progress on business and human rights more broadly.