

Watchdog Capture

The social licensing of child harm, a case study of Common Sense Media and OpenAI.

What happens when the watchdog partners with what it watches?

Common Sense Media is the child-safety backstop the public, press, and legislators rely on, **150M+ users, 1.4M educators**. In January 2024 it partnered with OpenAI. This is a structural account of what that partnership produced.

Threats to the public allege no individual wrongdoing. It names a power-concentration mechanism, *watchdog capture*, that operates below the threshold of formal regulation.

Apply the corrosive-capture lens to a watchdog, not a regulator.

CLASSICAL CAPTURE

A regulated industry comes to dominate its regulator. *(Stigler, Peltzman)*



CORROSIVE CAPTURE

Outputs diluted, slowed, restated, the **appearance** of oversight, minus its force. *(Carpenter & Moss)*



SOCIAL LICENSE

The community's permission to operate, manufactured by a trusted endorsement.

"...incredibly important for the societal license to be able to operate." , OpenAI chief global affairs officer, announcing the January 2026 ballot merger

A single-case study, built entirely on the public record.

What the record includes

CSM reports, Form 990 filings, joint press releases, the Raine complaint and OpenAI's answer, AB 1064, SB 243, the ballot measure text, verified news coverage.

What it excludes

Internal communications at either organization, board minutes, and privileged litigation material. No claim rests on unseen evidence.

The convention

Causal language follows litigation-reporting practice: users are described as dying after a pattern of use, not from a product, matching the litigation's own posture.

Journalist first. Litigation second. The watchdog third, every time.



SOURCE: TWENTY-THREE YEARS · ALWAYS THIRD. In each cycle, the watchdog then endorses the firm, on terms favoring the firm's remedy over its own earlier position.

Jan 2024	CSM & OpenAI announce their partnership.
Apr 11, 2025	Adam Raine, 16, dies by suicide after extended ChatGPT interaction. <i>(per complaint)</i>
Apr 30, 2025	CSM's first chatbot risk assessment, on companions. ChatGPT not a primary subject.
Aug 26, 2025	The Raine family files suit in San Francisco Superior Court.
Sep 2025	OpenAI ships parental controls; CSM calls them " <i>a good starting point</i> " that day.
Oct 13, 2025	Newsom vetoes CSM's AB 1064; signs the weaker SB 243.
Jan 2026	Merged "Parents and Kids Safe AI Act", drops the erotic-chatbot prohibition.

223

days

from Adam Raine's death to CSM's first risk assessment naming ChatGPT as a tested product.

A representative gap, not a unique one.

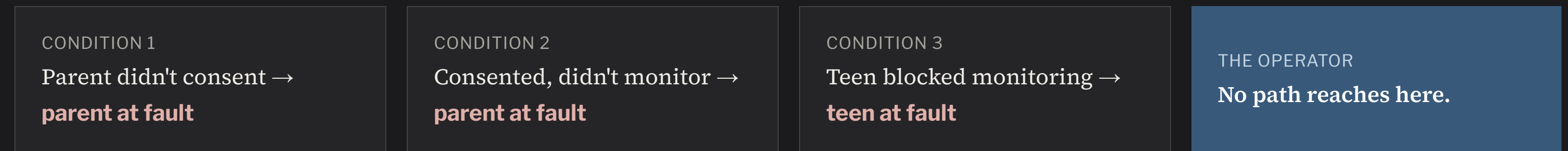
The April 30 companion-app assessment named ChatGPT only in a parenthetical. The first to test it came November 20, **86 days after the lawsuit was filed**. The observation rests not on what was known in April, but on the interval that followed.

"A good starting point."

The endorsement is not dishonest. But the controls are **opt-in for the teen, not the parent**, and a user told by the chatbot to hide the conversation has no incentive to link the account. The endorsement converts a liability-displacing feature into "a good starting point," supplying social license the architecture cannot.

SOURCE: THE NEW YORK TIMES. The watchdog's output is not eliminated. It is softened in a direction that materially benefits the firm.

Three exits. None of them terminates at the operator.



The same loop recurs in social-media age-gating, edtech consent flows, and fintech fraud reporting. The pattern generalizes.

AB 1064 · VETOED

Bar minors from companion chatbots

...unless the operator could demonstrate the chatbot could not produce sexual or self-harm content. **The watchdog's strongest position.**

→
DID NOT SURVIVE

MERGED MEASURE · JAN 2026

"Identify youth users; block manipulation & ads"

Prohibition replaced. Retains the three-hour disclosure interval, the firm's own competing-measure signature.

The brand changed. The product did not.

November 20, 2024: "ChatGPT Foundations for K-12 Educators," a free eight-lesson course, co-branded with the Common Sense logo and distributed at no cost to schools. The content is conversational guidance for teacher-led activities, not a design change to ChatGPT.

LEHANE'S OWN WORDS, A YEAR
LATER

The trust the partnership produced was “incredibly important for the societal license to be able to operate.”

Not a regulator captured, a watchdog. And it can be measured.

The firm acquires, through a private partnership, social license it could not credibly buy directly. The advocacy organization's reach is the resource. Three public-record indicators detect it:

01

Late-confirmation interval

Days from disclosure to guidance, the 223-day anchor.

02

Position-reversal record

Legislative or ballot positions surrendered after partnership.

03

Co-branded releases

The watchdog's logo on substantively unchanged practice.

One case. A general mechanism. It will keep producing the same outcomes until it is named.

The field has studied the model-builders, the compute, and the regulators. It has barely studied the validators whose endorsements **manufacture the social license** that lets firms keep operating without binding regulation.

— REFERENCES & SOURCES

Gilly, T. (2026). *Watchdog Capture and the Social Licensing of Child Harm: A Case Study of Common Sense Media and OpenAI*. Working paper (v4), Real Safety AI Foundation.

California Legislature (2025a) Assembly Bill 1064 (Bauer-Kahan), LEAD for Kids Act. Vetoed by Governor Newsom, 13 October. Sacramento, CA.	Common Sense Media (2025b) Common Sense Media finds major AI chatbots unsafe for teen mental health support, press release, 20 November. Available at: https://www.common sense media.org/press-releases/common-sense-media-finds-major-ai-chatbots-unsafe-for-teen-mental-health-support
California Legislature (2025b) Senate Bill 243 (Padilla), companion chatbots. Approved by Governor Newsom, 13 October. Sacramento, CA.	Common Sense Media (2025c) Big Tech lobby kills critical kids’ AI safety legislation, press release, 13 October. Available at: https://www.common sense media.org/press-releases/big-tech-lobby-kills-critical-k
Carpenter, D. and Moss, D. (2014) Preventing regulatory capture: special interest influence and how to limit it. Cambridge: Cambridge University Press.	Common Sense Media and OpenAI (2024) ‘Common Sense Media and OpenAI partner to help teens and families safely harness the potential of artificial intelligence’, press release, 29 January. Available at: https://www.common sense media.org/press-releases/common-sense-media-and-openai-partner-to-help-te <code>WatchdogCapture:CommonSenseMediaandOpenAI</code>
CNN (2025) ‘Kids and teens under 18 shouldn’t use AI companion apps, safety group says’, CNN, 30 April. Available at: https://www.cnn.com/2025/04/30/tech/ai-companion-chatbots-unsafe-for-kids-report	
Common Sense Media (2024) Annual report and program reach. San Francisco, CA: Common Sense Media.	
Common Sense Media (2025a) AI risk assessment: social AI companions. San Francisco, CA: Common Sense Media. Released 30 April. Available at: https://www.common sense media.org/sites/default/files/pug/csm-ai-risk-assessment-social-ai-companions_final.pdf	

— REFERENCES & SOURCES (CONTINUED)

Common Sense Media and OpenAI (2026) ‘Common Sense Media, OpenAI join forces on strongest youth AI safety measure in U.S.’, press release, 9 January. Available at: <https://www.common sense media.org/press-releases/common-sense-media-openai-join-forces-on-strongest-youth-ai-safety-me>

Gunningham, N., Kagan, R.A. and Thornton, D. (2004) ‘Social license and environmental protection: why businesses go beyond compliance’, *Law and Social Inquiry*, 29(2), pp. 307–341.

Hao, K. (2025) *Empire of AI: dreams and nightmares in Sam Altman’s OpenAI*. New York: Penguin Press.

Internal Revenue Service (2024) Common Sense Media Form 990, Fiscal Year 2023, EIN 41-2024986. Filed November 2024. Available at: <https://projects.propublica.org/nonprofits/organizations/412024986>

KQED (2026) Myrow, R. ‘OpenAI and Common Sense Media partner on new kids AI safety ballot measure’, *KQED News*, 9 January (updated 10 January). Available at: <https://www.kqed.org/news/12069286/openai-and-common-sense-media-partner-on-new-kids-ai-safety-ballot-measure>

Lessig, L. (2013) ‘Institutional corruption defined’, *Journal of Law, Medicine and Ethics*, 41(3), pp. 553–555.

Maheshwari, S. (2017) ‘On YouTube Kids, startling videos slip past filters’, *The New York Times*, 4 November.

OpenAI (2025) Introducing parental controls, press release, 29 September. Available at: <https://openai.com/index/introducing-parental-controls/>

OpenAI Answer (2025) Defendant OpenAI’s answer to plaintiffs’ complaint, *Raine v. OpenAI*, Case No. CGC-25- 628528, San Francisco Superior Court, filed 25 November.

OpenAI Help Center (2026) Parental controls FAQ. Available at: <https://help.openai.com/en/articles/10910070-parental-controls-faq>

Peltzman, S. (1976) ‘Toward a more general theory of regulation’, *Journal of Law and Economics*, 19(2), pp. 211–240.

Raine v. OpenAI (2025) Complaint, San Francisco Superior Court, Case No. CGC-25-628528, filed 26 August.

Stigler, G.J. (1971) ‘The theory of economic regulation’, *Bell Journal of Economics and Management Science*, 2(1), pp. 3–21.

TechPolicy Press (2025) ‘Breaking down the lawsuit against OpenAI over teen’s suicide’, *Tech Policy Press*, 26 August. Available at: <https://www.techpolicy.press/breaking-down-the-lawsuit-against-openai-over-t>

Verdegem, P. (2022) ‘Dismantling AI capitalism: the commons as an alternative to the power concentration of Big Tech’, *AI and Society*, 39(2), pp. 727–737. doi: 10.1007/s00146-022-01437-8.