AI DISCLOSURES PROJECT

OCTOBER 2025

POLICY BRIEF

Governing AI Through SEC Disclosure

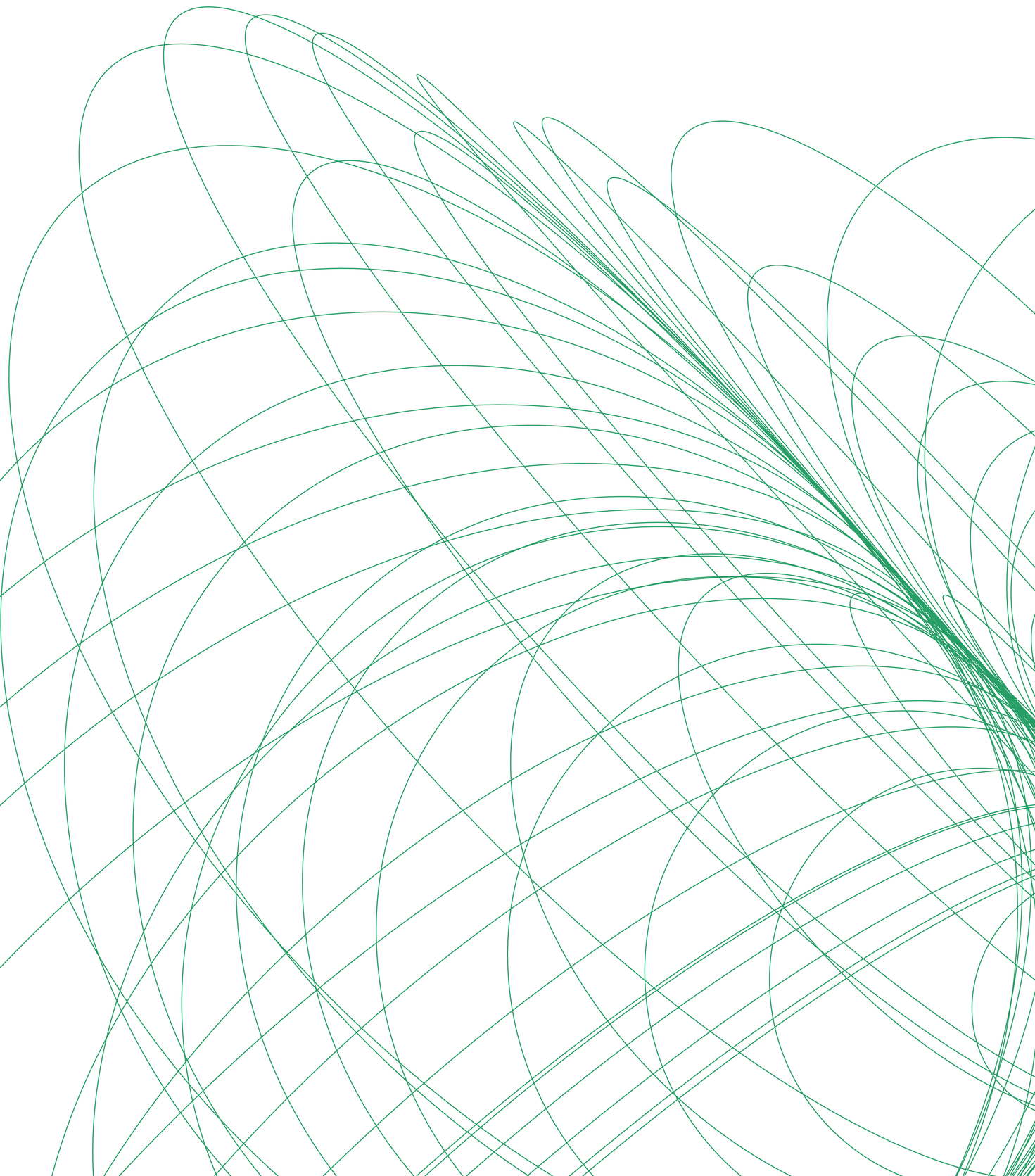
Materiality Standards and Incident Reporting—Lessons from Cybersecurity

Ilan Strauss
Program Director
Social Science Research Council

Tim O'Reilly
Program Director
Social Science Research Council

Sruly Rosenblat
Program Associate
Social Science Research Council

Isobel Moure
Program Associate
Social Science Research Council



100 YEARS | SOCIAL SCIENCE RESEARCH COUNCIL

About The AI Disclosures Project

Led by technologist Tim O’Reilly and economist Ilan Strauss, the AI Disclosures Project addresses the potentially harmful societal impacts of AI's unrestrained commercialization. By improving corporate and technological transparency and disclosure mechanisms, it aims to ensure that economic incentives don't compromise safety or equity, and avoid fostering excessive risks. Disclosures are vital for well-functioning markets yet remain lacking in AI. Just as financial disclosure standards fostered robust securities markets, standardized AI disclosures can build trust, expedite adoption, and spur innovation. Through research, collaboration, and policy engagement, the AI Disclosures Project aims to develop a systematic framework for meaningful “Generally Accepted AI Management Principles.” The project is generously funded by the Omidyar Network, Alfred P. Sloan Foundation, and Patrick J. McGovern Foundation.

DOI: 10.35650/AIDP.4120.d.2025

This policy brief can be referenced as follows:
Strauss, Ilan, O’Reilly, Tim, Rosenblat, Sruly, and Isobel Moure. “Governing AI Through SEC Disclosure.” *Policy Brief.* Social Science Research Council, October 2025. https://www.ssrc.org/publications/governing-ai-through-sec-disclosure-materiality-standards-and-incident-reporting-lessons-from-cybersecurity/

100 YEARS | SOCIAL SCIENCE RESEARCH COUNCIL

Website: AI Disclosures Project
Substack: Asimov’s Addendum
X: @AIDisclosures

Overview

AI technologies now contribute substantially to corporate performance and risk, but investors lack decision-useful information. Drawing on the SEC's 2023 cybersecurity model, we propose a materiality-first AI disclosure regime: SEC guidance clarifying what is material; a dedicated AI-incident item on Form 8-K; a standing 10-K section on AI strategy, governance, risk, and dependencies; and SEC enforcement against AI-washing. We urge reversing 2012 JOBS Act changes that let large private firms avoid public reporting and registration. Focusing on material impacts – not abstract capabilities – will discipline AI deployment and improve market oversight.

1. Introduction & Overview—The Need For Material AI Disclosures

The Trump administration's push to move away from quarterly corporate reporting has sparked a debate about the value of corporate reports, with important takeaways for AI-related reporting requirements.

John Authers noted in his recent column, quoting Sarah Williamson at FCLT Global, that disclosures are not really about timing, but "materiality" (i.e., importance): "What really matters…is the materiality of what to tell investors, not the periodicity." And the bar for what is material should be lower, she argues. That is, more things than are currently being disclosed should be considered materially important for companies to disclose, rather than fewer.

This principle has potentially far-reaching implications for AI disclosures. Rather than getting caught up in debates over how often AI companies should report, we should be asking: What risk events in AI systems are already material enough to investors to warrant immediate disclosure by corporations? And how exactly should these be disclosed by public companies?

In line with our previous work on corporate disclosures for large digital platforms with Prof. Mariana Mazzucato, we argue that disclosures must evolve with the changing structure of the economy, reflecting the new types of risks and operational facts that are material to investors. This ultimately prioritizes disclosure quality – judged by its relevance and depth – above reporting frequency.

Drawing on the SEC's 2023 rule on cybersecurity incident reporting, we propose the following steps to bring AI governance within the SEC's existing public reporting framework:

1. **Clarify how existing SEC disclosure rules apply to AI.** The SEC should issue "Disclosure Guidance" specifying what AI activities and risks are material and should be disclosed. Clearly defining material AI incidents – e.g., systematic model failures, major service outages, errors requiring widespread customer remediation, loss of essential third-party model access – will help companies disclose only what matters.

2. **Integrate AI-specific risks into existing disclosure filings.** In 2023, the SEC introduced a cyber rule requiring companies to report material cyber events within four days on Form 8-K and describe cyber-risk management in annual 10-Ks. Do the same for AI: add an AI incident item to the 8-K and require annual discussion of AI governance and risk management.

3. **Enforce the rules.** In crypto and cyber, improved disclosures followed real prosecutions. The SEC should continue to bring any material cases against AI washing, misleading claims, and fraud.

4. **Remove the loopholes that allow private companies to avoid going public.** For any of these disclosure obligations to apply to OpenAI and Anthropic, as private entities, we first need to reverse the 2012 JOBS Act changes that made it easier for large capital raisers to avoid public disclosure by raising money from private capital through special purpose vehicles (SPVs). Previously, companies like Google and Facebook became too big to remain private and were forced to go public. Those obligations were watered down significantly in 2012.

In contrast to managing AI through arbitrary technical thresholds, our approach emphasizes AI governance based initially on its potential material impacts to the real economy. The reporting metric is not an abstract AI capability but consequential effects on a company's operations, customers, or financial results – a language legislators, courts, investors, and the public already understand. Our recommendations support AI-related disclosure obligations that are more comparable, timely, and granular — and that incentivize greater company risk mitigation measures.

2. AI as a Market Technology—Materiality and Disclosure

Paul Atkins, the chair of the U.S. Securities and Exchange Commission (SEC), argues that corporate disclosures should be driven by the materiality principle – disclose what a reasonable investor would care about when making an investment decision. Let's "stick to business." **And what is becoming a bigger business than AI?**

AI's contribution to the economy is already staggering: In the first half of 2025, AI-related capital expenditures would have contributed 1.1% to GDP growth, more than the U.S. consumer, if the capital items were not largely imported. As a percent of GDP, capital expenditures on data centers (1.2%) were greater than the telecom spending in all of 2020 (1%), estimates economist Paul Kedrosky. However, much of these are imported capital inputs and so would be deductions from GDP. Investments in AI are on track to surpass those made in the internet during the boom years of 1995-2000. Meanwhile 80% of the stock market gains in 2025 (until October) were due to AI companies.

Princeton computer scientists Arvind Narayanan and Sayash Kapoor call **AI a "normal technology"**—transformative but not unlike previous inventions, such as railroads or electricity, where impacts were felt gradually over time as adoption ramped up. ChatGPT's rapid uptake illustrates that digital markets are the ultimate 'normalizing' force. But once a market takes hold, its logic imprints itself into a technology's DNA. Social media began as a way to stay connected with friends, but monetization pressures transformed it into an engagement-maximizing machine – an endless scroll designed to keep us hooked. OpenAI CEO Sam Altman calls algorithmic feeds "the first at-scale misaligned AIs." AI's sycophantic capabilities, monetized as companions or bottomless video feeds, exhibit a similar trajectory.

Given AI's commercial character, public oversight should start with the SEC's corporate disclosure regime. AI markets currently lack full and timely information, since prominent AI companies remain private and, in the absence of guidance, companies disclose platitudes. In turn, allocations of AI-related capital cannot be properly evaluated, litigation is ballooning, 'AI washing' and fraud are commonplace, and technologies are being deployed prematurely under a 'move fast and break things' ethos.

After the 1929 crash, the SEC mandated corporate disclosures to surface material risks to investors by requiring companies to publish annual 10-K reports, quarterly 10-Qs, and event-driven 8-Ks when an incident occurs. This remains one of the few proven systems for assessing corporate risk at scale.

The SEC's 'materiality' standard transforms private knowledge into public disclosure, creating the information substrate on which markets for audit, insurance, and research can operate. This ecosystem doesn't just inform, though – it disciplines: rewarding sound AI governance through lower capital costs and punishing poor risk management through market pressures.

Unfortunately, OpenAI and Anthropic are private, so they sit outside the SEC's public-company disclosure regime, despite being multibillion-dollar enterprises. (Private companies remain subject to SEC anti-fraud rules and investor protections, but are exempt from ongoing disclosure requirements.) Previously, companies like Google and Facebook were compelled to go public as they grew up. But the 2012 JOBS Act raised the threshold for mandatory registration fourfold and loosened restrictions on private fundraising. Capital raised through special purpose vehicles (SPVs) count as a single shareholder. And the threshold of total shareholders needed to be met in order to force public registration was raised from 500 shareholders of record to 2,000 shareholders of record – or 500 non-accredited investors (whichever comes first). Employees who received stock compensation are excluded from the count entirely. Most venture investors and many employees are accredited investors (high income / net worth), and so do not count toward the 500 limit anymore.

The above 2012 JOBS Act changes – combined with the incredible growth in VC, private equity, and sovereign wealth funds capital – has resulted in an explosion of late-stage private capital raising that enables very large companies to stay private and avoid public reporting requirements. The role of venture private capital in funding private AI companies is unprecedented (slide 22) compared with other technologies. In 2025, OpenAI, Anthropic, and xAI "captured" over $50 billion in VC funding. OpenAI's capped-profit partnership and Anthropic's public-benefit corporation might sound civic-minded, but ultimately insulate them from market-oversight.

Despite being private, OpenAI is causing major swings in the public stock market. OpenAI's decision to partner with Shopify, Etsy, and now Advanced Micro Devices (AMD) has sent their shares soaring. Its partnerships with chipmakers AMD and Nvidia involve fairly opaque, "circular", financing deals. Notes a leading investment strategist in Bloomberg: "it is certainly an odd situation for a private company to have so much impact…[OpenAI] can be more agile and creative, and that leads to the ripple effect we see in other companies, both good and bad." Further complicating matters is the incredible investments made in private AI companies by major listed companies, such as Amazon in Anthropic and Microsoft in OpenAI. This increases risks arising from the reverse impacts of these investments on to the investing company.

3. Cyber Risks as a Model Disclosure Framework for AI-Related Risks?

Perhaps as a result of a recent proposal from the Long Term Stock Exchange (Disclosure: Tim O'Reilly is an investor), which was reported on by *The Wall Street Journal* on September 8, President Trump proposed that public companies' quarterly reporting should instead become bi-annual (twice a year). Trump certainly made the case for it based on the same LTSE argument: that quarterly reporting places undue burdens on public companies and pushes executives into short-termism – so-called "expectations management."

Time-based reporting requirements incentivize companies to structure decisions around a company's financial calendar. But this might delay crucial information from being released by companies to the public as they occur.

Enter the 8-K Form. The 8-K can be thought of as a breaking news bulletin, used by the corporations to announce significant events within four business days. The list of "Items" that triggers a filing can vary. Many items are triggered automatically by an event, like a corporate bankruptcy. Other items on the list are based on the company's judgment around whether the event is "material"[2] i.e., sufficiently likely that a reasonable investor would care about it – and only then would they file an 8-K form. A cybersecurity incident is one such "if it's important enough" thing to disclose (Item 1.05).

[2] **"Material"** is defined by the courts to mean whether there is a substantial likelihood a reasonable investor would view it as important (TSC v. Northway), and weighs the probability of an event happening against its magnitude (Basic v. Levinson) – an approach the AI risk community also buys into.

An important and relatively new corporate disclosure requirement **that uses the 8-K Form is the SEC's 2023 Cybersecurity Incident rule for public companies.** The Cyber rule says that when a company suffers a material cybersecurity incident it must report it to shareholders within four business days through the 8-K Form. And when it's a material cyber event impacting shareholders, then it can be filed through the newly added Item 1.05, **specifically for cyber incidents.**

In combination with strong SEC enforcement, the Cyber rule seems to have worked. Cyber incidents are disclosed in a far more timely and comparable manner now, and companies appear to be devoting more resources to the problem. *Moreover, companies absorbed these new requirements with ease because they were well prepared from previous guidance.*

Part of the Rule's innovation is that the material event-triggered 8-K filing for cyber incidents sits **alongside a standing annual 10-K disclosure requirement for companies specifically for cyber-related issues (Reg S-K Item 106)**, covering things like board and management oversight, processes for identifying and managing material cyber risks, whether such risks materially affect the company, and more.

The question then is whether AI-specific risks require a similar treatment to cyber ones. Below we show that a substantial "disclosure gap" already exists for AI. This is the gap between the AI-risks already out there facing AI companies, and what they are currently disclosing.

4. Evidence on the Disclosure Gap

What disclosure gaps exist? Put differently: what reporting would enable responsible investment decisions about AI companies? Share prices can't reveal inadequate disclosure – since by definition they only incorporate information already available to the market.

4.1 Litigation

Litigation shows market dissatisfaction with existing corporate AI disclosures to be high. Fisher Phillips' AI litigation tracker for the U.S. currently shows 92 cases.[3] Litigation on securities class action lawsuits covering false or misleading statements on AI is on a near exponential rise in the U.S., from 7 cases in 2023, 14 cases in 2024, and 12 cases so far in 2025. AI-usage now exposes companies to a range of risks from product liability & negligence, wrongful death, defamation, and publicity and privacy, to name but a few. For example, in Garcia v. Character.AI & Google, the court has let the case proceed (May 22, 2025) over a teen's suicide, allegedly encouraged by a chatbot's messages. Claims include wrongful death, negligence, and deceptive trade practices.

4.2 10-K Disclosures

To manage growing AI-specific risks, companies are signaling greater disclosure to shareholders, but only superficially. An analysis by Arize AI, as reported by the *Financial Times*, found that 56% of Fortune 500 companies cited AI as a "risk factor" in their most recent 2024 annual 10-K reports.[4] Netflix, Motorola, and Salesforce all discuss AI-specific risks – yet only in superficial boilerplate terms, according to a recent and comprehensive academic study on 10-K disclosures for AI. Similarly, SEC staff letters to companies show that much of the guidance was thin on details. Staff consistently requested more specifics from companies on disclosure details for AI-related topics.

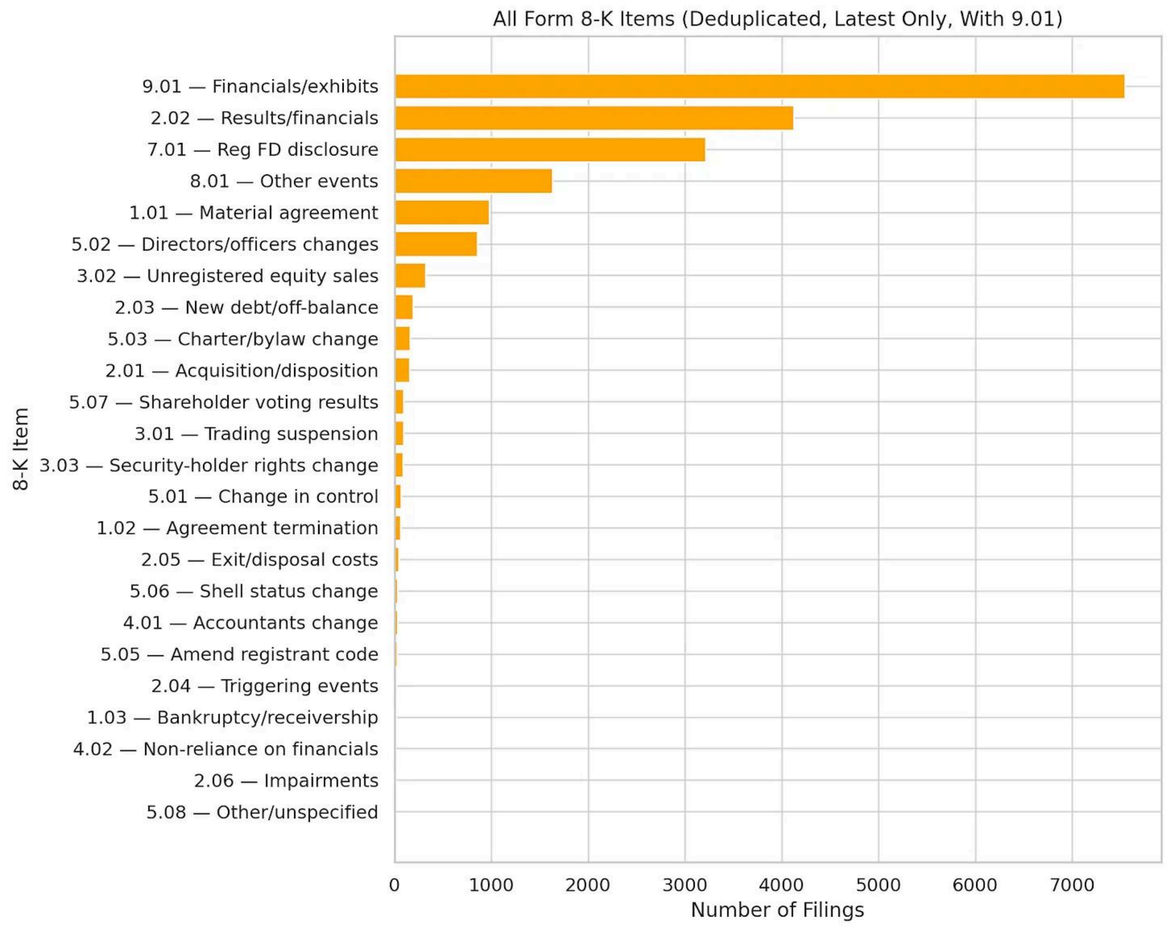
Aware of the AI-disclosures gap, the SEC launched in 2024 AI-specific guidance and enforcement covering AI washing, conflicts of interest, and systemic risk, along with enforcement actions.[5] The SEC now even has a newly dedicated Chief AI Officer (CAIO) Valerie A. Szczepanik, who will oversee a new SEC AI Task Force, though its focus is more on internal innovations.

4.3 8-K Disclosures

To analyze 8-K filings, we constructed our own dataset of all AI-related event-driven filings between November 1, 2022 until September 18, 2025, covering 1,741 corporate issuers. It highlights at least three key AI-related corporate disclosure gaps:

No Risks Here. The first is that 8-K disclosures almost exclusively concern a company's commercial ventures (Figure 1 below), covering important agreements (Item 1.01), such as model licensing, cloud/compute commitments, strategic data deals, and reseller/partnership agreements; but also Financial matters (Item 2.02).

Figure 1. 8-K filings by public companies in the U.S. on artificial intelligence and generative AI, by Item topic.



Note: n = 7,856 (deduplicated). November 1, 2022 until September 18, 2025. See EDGAR: https://www.sec.gov/edgar/search/efts-faq.html

[4] Alston & Bird's 2024 study found that 46% of Fortune 100 companies included AI-related risk disclosures in their annual 10-K forms. Disclosures fell broadly into five buckets: (1) cybersecurity risk; (2) regulatory risk; (3) ethical and reputational risk; (4) operational risk; and (5) competition risk.
[5] The conflict of interest proposal was withdrawn under the Trump administration.

Safety and guardrails, i.e., AI-risks, rarely feature. ***Overwhelmingly most AI filings at 66% (two-thirds) are positive in nature.*** In other words, companies have a tendency to use the 8-K to alert investors to news that may help their business prospects.

Using GPT 5 nano, we classified filings containing sufficient text into three buckets of sentiment and found 4,952 8-K filings with "positive" sentiment, 1,367 as "negative", and 1,121 as neutral. We were unable to extract text from all disclosure filings.

Confusion reigns. Most disclosures for AI-related impacts are through Item 8.01: a voluntary catch-all event category useful for AI updates that are not yet a mandated material trigger but still market-relevant. This implies that companies are not yet sure where to put such AI-triggered events – or are unsure when an event is sufficiently material to disclose it elsewhere.

Big firms need a 10-K mandate. Finally, 8-K filings on AI-related matters are driven by smaller companies since submissions reflect the universe of filing firms. Big Tech's 8-K disclosures are not very prominent – as expected – since they only constitute 0.34% of the companies making 8-K submissions in our data. AMZN made 14 submissions, followed by NVDA (11), MSFT (10), and META (12), and GOOGL/GOOG (9 each).

Practically, this means that any new 10-K requirement that covers AI-specific business activities and risks in detail could significantly enhance market transparency, since these mega-cap firms have an outsized impact on the AI market (together with OpenAI, Anthropic, and a few others).

5. So What Should We Be Aiming For?

I) SEC Guidance Note on AI. To get the ball rolling, an SEC guidance note (called "CF Disclosure Guidance") could help companies understand how existing company disclosure rules apply to AI-related matters. *At its core this should define material AI incidents in plain English* to include systemic model failures, major outages, widespread customer remediation, loss of essential third-party model access, impactful changes to safety guardrails, and so on. *It should also clarify how AI-driven events fit within existing 8-K categories and how to disclose relevant AI-activities within a company's annual 10-K report.*

A guidance note is not binding law, but it can strongly influence company filings and SEC actions. *For example, the 2011 Cybersecurity memo (Topic No. 2) told issuers what to discuss under Risk Factors, MD&A, Business, and other items in their 10-K report.*

An AI guidance note would provide the same practical roadmap as the 2011 Cyber memo: specific, concrete examples showing companies how to disclose AI risks and opportunities substantively across business operations (S-K Item 101), risk factors (Item 105), trends and uncertainties in MD&A (Item 303), and other key sections – avoiding generic boilerplate.

Part of the guidance might encompass **AI-related escalation criteria for potential Form 8-K reporting.** Companies should maintain disclosure controls that identify AI-related developments which, **if material,** may require a current report on Form 8-K.

Quantitative indicators (escalation): Statistically significant deviations from historical baselines in KPIs plausibly affected by AI system changes – e.g., engagement (DAU/MAU, time-on-platform), monetization (CTR, conversions), and risk metrics (credit approval or denial rates, charge-offs, loss ratios, harmful-output and jailbreak rates, fraud-detection efficacy). Indicators inform but do not by themselves determine materiality and Item applicability, consistent with SEC KPI/MD&A guidance.

Qualitative indicators (escalation): Changes to AI objectives, guardrails and policies with expected impact on harmful-output rates or regulatory exposure; material data-provenance shifts (e.g., addition of sensitive datasets); dependency changes (e.g., migration of core functionality to third-party models and APIs); or significant compute capacity loss or outage. Counsel should assess whether any specific 8-K Item is implicated (e.g., Items 1.05, 1.01, 2.06, or 8.01).

II) **Create a new AI-risk item on the 8-K disclosure Form for material AI-driven events as they happen** – *modeled on Item 1.05, Cybersecurity Incidents (2023)*. Note that it is not the technology (AI) itself that triggers a filing rather than a material, incident-style impact. The trigger is not "an AI model changed," but that "the change or failure had a meaningful effect on operations, customers, compliance, or financial results."

Companies already use the 8-K Form to alert investors when something important happens between annual or quarterly reports. The idea here is to add a dedicated item for ***AI-related material incidents***, so that there is a clear place to report them when they matter. This can help ensure that companies do not skip reporting the "risks" when disclosing material AI-related events.

An "AI incident" is a development arising from the use of AI systems that has a meaningful effect on the business. Examples include: a model failure that misprices loans; an AI system outage interrupting service; an AI-driven error requiring customer remediation; or a sudden loss of access to a third-party model on which a product depends. The trigger is the impact itself.

III) *Add a standing AI section in the annual 10-K Form that explains how a company manages AI.* One-off 8-K event reports are, by themselves, insufficient. Investors also need a clear, yearly picture of how a company runs its AI-related activities, covering: how it is used in products and operations, who oversees it, what the main risks are, and what controls are in place. A new 10-K item would provide that exact structure, thereby encouraging companies themselves to adopt a longer view of these risks.

Companies would explain their approach to risk management (how they test and monitor systems, how they roll out changes, how they respond when something goes wrong); their strategy (where AI fits in the business and why); and their governance (who is accountable at the management and board level). They would also describe key dependencies that could affect reliability or cost (such as reliance on outside model providers, critical data sources, or a single cloud vendor), along with any concentration risks that come with those choices.

The goal is not to jam in unnecessary detail into the 10-K but to make the business implications of AI understandable to the investing public: where the leverage points are, how failure is prevented, and what the plan is when problems occur.

Finally, labeling the main AI elements with standard, machine-readable (iXBRL) tags (the same way the SEC does for several other disclosures, such as the SEC's cyber rule) would let analysts and watchdogs compare companies more easily and spot patterns over time.

IV) ***Enforce the rules.*** In crypto and cyber, improved disclosures followed real prosecutions. The SEC should continue to bring any material cases against AI washing, misleading claims, and fraud.

V) **Reverse the JOBS Act loopholes** that allow companies to raise billions from hundreds of investors while remaining private. If you access public savings at scale, you should meet public disclosure standards. So-called "Regulation D" exemptions[6] currently permit unlimited private capital raises from accredited investors by a company without triggering reporting requirements. We propose: treating SPVs as look-through entities so consolidated shareholder counts cannot be gamed; narrowing the employee shareholder exemption; and capping Reg D fundraising (e.g., $1 billion lifetime or 250 shareholders) before companies must register as reporting entities.

6. Conclusion: Empower Markets with Information

Unlike capability thresholds, the SEC disclosure approach for public reporting companies anchors oversight in materiality: what AI does to a firm's operations, customers, and earnings that an investor would care about. It rewards evidence – not hype. It is a language investors, courts, and boards already understand.

No disclosure regime will fix every AI risk. But a materiality-based framework can better align company incentives, surface urgent hazards, and give democratic institutions leverage over a profoundly commercial technology. If quarterly reporting goes, the quid pro quo should be stronger event-driven transparency and annual reporting.

---

[6] Regulation D provides exemptions from SEC registration requirements, allowing companies to raise unlimited capital from accredited investors without: Registering the securities offering with the SEC; and Becoming a public reporting company (filing 10-Ks, 10-Qs, etc.). The key rules were Rule 506(b): Raise unlimited amounts from accredited investors (and up to 35 sophisticated non-accredited investors) without general solicitation; and Rule 506(c): (Added by JOBS Act) Same, but allows public advertising – only to verified accredited investors