AI Disclosures Project

Building an Architecture of Participation for Human+AI Markets in 2026 & Beyond

Tim O'Reilly and Dr. Ilan Strauss

A Report to Donors on the Bellagio *Human+AI Markets Convening* (April 27–May 1, 2026) and the Future Work Program of the AI Disclosures Project (CS&S)

Table of Contents

AI Disclosures Project | Hosted at Code for Science and Society (CS&S): a registered US 501(c)(3) nonprofit
Collider, 26 Broadway, Floor 3, New York, NY 10004
https://ai-disclosures.org/

1. Mechanisms & Architectures for the Human+AI Economy

The welfare impact of AI on society will be a function of how consciously we design robust and participatory Human+AI markets: whose *mechanisms* embed human participation, whose *distributed designs* anticipate failure modes, and whose open, modular *architectures* – rather than closed, monolithic ones – let the market evolve in ways that resist single-gatekeeper capture. In this "market-architecture" view, who wins and who loses from AI will be shaped by our ability to overcome the familiar pathologies of closed vs. open architectures and code; to counteract the concentration pressures arising from AI models' high fixed training cost and variable inference cost; and to design open infrastructure capable of advancing user-centric data- and context-rich AI application markets.

Open protocols[1] – permissionless, modular rules for communication and interoperation – are a concrete response to the above challenges, providing new participants with an operational basis on which to build more participatory digital market infrastructure. AI's protocol layers (e.g., MCP, ACP) and standards (**skills.md**) work to structure AI markets in a modular, permissionless fashion, much as TCP/IP and HTTP did in the prior web era. But for open protocols to reach more deeply into the internet's evolving application layers, including how AI's inputs and outputs allocate value, associated economic mechanisms, licensing and standards bodies, and careful architectural design will be required to tip markets toward open, broad-based value creation and distribution.

The Rockefeller-sponsored Bellagio Convening (April 27 – May 1, 2026) was our first major gathering organized around the thesis that AI markets must be consciously designed for the welfare outcomes and architectural properties we want, drawing on a mix of experts in systems architecture, protocol design, economic mechanism design, and political economy.

2. The Bellagio Convening

"**Missing Mechanisms and Architectures for the Future Human+AI Economy**" brought together twenty leaders[2] hosted at the Rockefeller Foundation's Bellagio Center from April 27 to May 1, 2026.

[1] A protocol is a specification of how independent systems communicate or interoperate. Open protocols let independently built implementations work together without anyone needing permission from a central authority. "Open" implies that the spec is publicly documented, anyone can implement it without licensing, and it's typically maintained by a neutral body (IETF, W3C) rather than a single firm.

[2] Two participants fell sick en route to the convening, Andrew Trask (OpenMined and Google DeepMind) and Sarah Pearson (Creative Commons). We were fortunate enough to have Markus Mobius (Principal Researcher, Microsoft Research) join us last minute to help take one of the vacant slots.

Inspired by Tim O'Reilly's original 2004 conceptualization of the web's unique architecture of participation for users and not just software developers – enabled by open "communications protocols", "modular architecture", "hyperlinking", and other mechanisms for contribution – AI markets should be built with this end in mind too (read more: here and here). Facilitating inclusive AI markets at a more granular level will require addressing *The Missing Mechanisms of the Agentic Economy*, which include rules and institutions to align often conflicting incentives around participatory, distributed market outcomes.

The final Bellagio convening programme can be found here, along with participant bios, and background reading materials that were made accessible via our own MCP server and LLM model. Most of the 20 or so lightning-talk presentations can be found here (Day 2), along with minutes from the working groups (Day 3).

Who was in the room

A unique group of cross-disciplinary experts were assembled:

- **Co-hosts and Organizing Team:** *Tim O'Reilly* (CEO O'Reilly Media, AI Disclosures Project), *Ilan Strauss* (AI Disclosures Project, University College London), and *Sruly Rosenblat* (AI Disclosures Project).
- **Mechanism design and platform economics:** *Nicole Immorlica* (Yale Professor of Computer Science, Microsoft Research), *Markus Mobius* (Principal, Microsoft Research), and *Jacques Crémer* (Professor, Toulouse School of Economics; lead author of the European Commission's *Competition Policy for the Digital Era* report).
- **Web standards and internet architecture:** *Mark Nottingham* (Cloudflare Standards Lead, IETF Signals Working Group Head), *Mike Linksvayer* (VP, GitHub), *Alissa Cooper* (Head, Knight-Georgetown Institute), and *Mallory Knodel* (Co-Founder, Social Web Foundation).
- **Agentic infrastructure and protocol builders:** *Alex Hancock* (*goose* / Block – Agentic AI Foundation), *Liad Yosef* (MCP apps, Sequoia-backed AI startup), *Ido Salomon* (MCP apps, Monday.com AI Lead), *Ola Hungerford* (MCP + agent skills, Nordstrom), and *Tadas Antanavicius* (MCPulse).
- **Attribution and Publishing Markets**. *Geoffrey Bilder* (Chief Technologist, MIT Press), *Lucky Gunasekara* (CEO, Miso.ai), and *Paul Farrow* (Monetization Lead, Microsoft Content Marketplace).
- **LLM Model Expert:** *Serkan Piantino* (ex-Facebook FAIR, ex-VP Product Reddit); and an MIT scholar with almost two decades of experience in training language models who requested to remain *anonymous*.

The convening included: discussion groups on value and scarcity in the age of AI, dedicated time to expanding connections and inter-disciplinary thinking, twenty participant presentations to share

their expertise – with special confidential presentations provided by Paul Farrow on Microsoft's Content Marketplace (the only fully integrated programmatic AI content marketplace), and by an MIT scholar on how LLMs are trained, and how they value and store information, grounded in the foundational concepts of Shannon, Turing, and I.J. Good.

Three days of focus on market design and protocols

The three day schedule was designed to progressively move the room from theory to specification to tangible work programs:

Day 1 — Fundamentals (the group, value, mechanisms, entropy). Tim O'Reilly opened with why "mechanism design" is the right frame for the agentic economy – akin to Hal Varian's market shaping role at early Google. Ilan led a structured breakout on what becomes scarce and what becomes commoditized under AI. Nicole Immorlica provided a mechanism design tutorial covering canonical market models, signaling and information design (lemons), and Pandora's-box-style consumer search with the Weitzman index — work that has direct application to how AI agents will discover content and clear markets. Paul Farrow led a confidential discussion of Microsoft's publisher content marketplace for AI answer engines. The day closed with a technical presentation by an MIT scholar on the utility of information represented inside a large language model's weights (with implications for how data is stored and valued by LLMs).

Day 2 — Going Deeper (expertise). Day 2 allowed for deep dives into each person's area of expertise, providing a forward-looking view of where AI markets and AI infrastructure are heading (presentations here). We arranged the presentations into four broad categories: (1) "Agentic and AI Architectures", (2) "Web Architectures under AI", (3) "Model Context Protocol (MCP) and emerging AI protocols" and (4) "Agent Skills".

For "*agentic and AI architectures*": Nicole Immorlica and Markus Mobius painted an 'Me Agent' future where agents communicate on behalf of their humans in trusted enclaves. The enclave allows for rich private data to be shared between each person's agents while preserving the data's privacy from the other human. Mallory Knodel discussed the limits of end-to-end encryption for AI features in WhatsApp and other communication applications; Lucky Gunasekara addressed the challenges publishers face from scrapers, and what can be done about them; and Alex Hancock discussed goose as an open-source agent host, and how it can help shape ecosystem standards.

For "web architectures under AI": Mark Nottingham previewed competing IETF proposals to facilitate payment, agent communication and bot auth on the web; Mike Linksvayer discussed a "stamps" layer as a way to show payment and grant access to content, and how such a system complements and mirrors HTTPS; Liad Yosef discussed the 'last mile' problem of AI, where the human may need to take over from the model at some point and how MCP Apps / UI helps with this through integrating a company's existing UI into an LLM interface; and Alissa Cooper pitched a neutral web index, as a blank slate canvas idea, with the goal of alleviating the strain put on content creators by unauthorized scrapers.

We then went deeper into MCP: Alex Hancock demoed an agentic payment system using MCP and x402 protocol; Ido Salomon discussed how discovery of MCP apps currently happens and what the future of discovery looks like; Ola Hungerford explored the proposed skill extension to MCP, and how interceptors may soon be integrated into MCP – allowing for input to tools to be validated prior to sending to the MCP server and for outputs to be inspected on the way back.

Lastly, we went further into agent skills and the surrounding ecosystem. Tadas Antanavicius discussed how skills, MCP, and hooks can be bundled into sharable artifacts to use throughout an organization and showed what he is building for companies to facilitate this; Sruly Rosenblat discussed the difficulties in differentiating between what a skill actually covers relative to reference documentation / set of capabilities, and demoed a system to evaluate this; and Lucky Gunasekara with Tim O'Reilly discussed challenges in creating author skills on the O'Reilly Media platform.

Day 3 — Prototyping Working Groups. We started off with Mark Nottingham presenting on the importance of independent user-agents and how power dynamics are mediated in internet architectures. Afterwards, three working groups were established to work on topics of most mutual interest and promise for future prototyping.

Agent me. *The first group focused on agent-to-agent communication and how to build safe and private communication that acted in the interest of the principal user*. Continuing on from the *Me Agent* idea, proposed by Nicole Immorlica and Markus Mobius from Microsoft, a trusted execution environment (TEE) was proposed alongside a 'progressive disclosure' scheme where agents only reveal the information from their human needed to receive an answer in return.

Agent telemetry. *A second group focused on ensuring that servers / tool owners have access to user telemetry on how their data is consumed, as mediated by the agent*. The problem is how to assess a user's intent in a world where people are not clicking links and how data on the users' intent could be sent to the companies who provide a given information source or tool. An official MCP extension was proposed where agents would be able to send their reasoning explaining why they requested specific information or used a specific tool alongside their tool call.

A neutral web index. *The third (and final) group discussed how to conceptualize and create a viable alternative market to the existing grey market where LLMs scrape for content*. A "neutral web index" or repository was proposed, based on indexing and recording an item's informational uniqueness (a patent / priority fingerprint of sorts) relative to other information, which no individual company can reproduce through scraping. The index would show how informational records relate to one another through differences in their entropy and other relevant statistics. Monetization can be built on top of this. A finger printing prototype has already been constructed.

What Bellagio delivered

1. **A shared map.** Twenty experts from very different vantage points – economists, protocol architects, lab policy leads, artists' advocates – found they had been drawing the same map in different vocabularies. A shared recognition emerged that Human+AI markets need

to be actively designed through its architecture and mechanisms (that shape institutions and rules), and that open protocols are a very concrete way to do so.

2. **Ongoing collaborations and working specifications.** Ongoing working groups and collaborations, some of which will convene post-Bellagio:

    a. A post-Bellagio convening note on the power of architecture and mechanisms to shape AI markets, led by the AI Disclosures Project and Cloudflare/IETF (Mark Nottingham).

    b. Agent communications over federated protocols, currently designed primarily for human communication, such as ActivityPub and Nostr, using *broadcasting* as a central feature to make discovery native to the protocol itself. This will aim to replace the current status quo of one-to-one agent communication which contains little to no agent discovery mechanisms, while simultaneously opening the door to humans and agents sharing one communication platform. When advancing it beyond the prototype stage, we will seek partnership and feedback from the W3C ActivityPub community, the Nostr ecosystem, the Linux Foundation's Agent2Agent (A2A) project, AGNTCY (Cisco's Internet of Agents initiative), and Mark Nottingham at the IETF.

    c. Building out an open source memory architecture to break down existing siloed and bundled memory systems (what we call the "memory walled garden"). Our prototype MCP server demonstrates cross-platform memory sharing, through a standardized interoperable and modular memory system, implemented through MCP transport. Leveraging existing collaborations with the Goose team and the Agentic AI Foundation, to bring full cross-platform chat history to coding agents, is the goal.

    d. A “neutral web index” (working group 3), as an alternative to web scraping, is continuing its work and applying to the Internet Society Foundation for funding, led by the AI Disclosures Project and supported by the Knight-Georgetown Institute and Creative Commons (partnership in discussion);

    e. A note on “mechanism design for AI markets” written by Jacques Crémer.

3. **A theory of change: protocols and power for modular architectures.** Mark Nottingham advanced the W3C priority of constituencies: users above publishers, above implementers, above specifiers, above theoretical purists, into the AI conversation as a concrete tool for adjudicating design trade-offs. Tim O'Reilly connected this to his longer-running "architecture of participation" argument (after Kapor and Lessig's “architecture is politics”): where modular, protocol-based systems such as Unix, the Internet, and the Web outscale monolithic ones, because their architecture turns self-interested user activity into collective value as a byproduct, rather than concentrating it in a single gatekeeper. MCP is the latest test of that proposition. Incumbents currently do

have an incentive to expose their data and capabilities through user-friendly MCPs, so the leverage sits with disruptors and with groups organizing around open standards and implementations to make the most of this momentary alignment of incentives. “Rough consensus and running code,” in the IETF tradition, was adopted as the operating principle.

This will be written down in a post-convening note that articulates a theory of digital market design stemming from architecture on the one end and mechanisms on the other.

Select Participant Feedback

“I am very grateful to have been included in the convening. Everyone in attendance was smart and came with an open mind to try to help. At the beginning of every technological revolution decisions that seem small or isolated at the time can have major effects on the years-long time scale. I hope that some of the work we began in Bellagio can help the AI economy develop in a more fair and equitable way than it may have otherwise”

– *Alexander Hancock* (goose lead – Block, Agentic AI Foundation)

“The Bellagio convening was, on every measure that matters to me, hugely positive. Its principal value lay in the composition of the participants. Sustained dialogue across the range of perspectives represented is rare, and the format permitted disagreement of a kind that is difficult to engineer elsewhere. I have continued to draw on several of the relationships formed there, and the discussions have informed my view of which problems in the emerging AI content economy warrant serious attention.

The substantive work on standards was, for me, the second significant outcome. Questions of disclosure, attribution and the protocols governing AI-publisher relationships are ones I had previously considered largely from within a single commercial vantage point. The opportunity to examine them alongside others approaching the same questions from different positions materially altered how I frame them.

The most concrete consequence on my return to Microsoft has been internal. I have argued, in a number of forums, for deeper calibration between Microsoft colleagues and the work of standards creation and standards setting, in the interests of the wider economy rather than any single participant in it. The convening reinforced for me how much rests on getting that engagement right.”

– *Dr Paul Farrow* (Principal Product Manager, Microsoft Advertising)

3. Way Forward

The agentic economy is being built right now, and largely without many of the architectures and mechanisms that prevent winner-take-most outcomes. Agents currently transact on users' behalf without disclosing whose interests they serve. Skill marketplaces host hundreds of malicious entries because discovery is unsolved. The web's crawler economy is collapsing under bot traffic nobody priced. The design choices being made today, and our ability to integrate open protocols, open architectures, and open mechanisms into AI markets over the next eighteen months will determine its evolving structure for years to come.

The AI Disclosures Project's 2026-2027 program, fiscally hosted now at Code for Science & Society, is a focused bet that the place to intervene is the protocol layer, not the regulatory layer that arrives after the fact. We are seeking $750,000 in a second round of funding. This builds on our initial round of generous funding support from the Rockefeller Foundation (for Bellagio / convening-only support), the Alfred P. Sloan Foundation, the Patrick J. McGovern Foundation, and the Omidyar Network, with seed funding provided by the O'Reilly Foundation. The AI Disclosures Project's 2026-2027 program is organized around three reinforcing approaches to shaping and realizing pro-social AI market outcomes:

Our Approach

Mechanisms. Under private information and strategic behaviour, what rules of the game and institutions produce a participatory AI economy rather than a brokered one? This is mechanism design in the Hurwicz–Myerson–Maskin tradition, applied to a market where humans, agents, and models all transact on each other's behalf, and where the ‘outside option’ (scraping or exclusionary deals) is appealing.

Architectures. What protocols and host applications make those mechanisms more economical and beneficial to deploy? Architecture is where market design gets enforced through capability rather than adjudication. MCP has already shown that an open standard can decouple tooling and even context from platforms; the question now is how to cement this initial victory.

Disclosures. Of the information that AI systems and markets generate and rely on, what is made visible, to whom, at what cost, and on what authority? A market without disclosure is one we cannot inspect, and an agent without disclosure is one we cannot trust.

The convenings, deliverables, and partnerships below convert this agenda into a 12–18 month operating plan.

Convenings: Designing the AI economy in public

The convenings are where the AI Disclosures Project puts mechanism designers, protocol architects, AI lab policy leads, and artist-side advocates in one room and forces an agenda to emerge.

1. **Bellagio Center (April 2026) — complete**. Twenty participants across the IETF, MIT CSAIL, Yale, Microsoft Research, Block, Cloudflare, and Knight-Georgetown Institute. The output of Bellagio is the program described in this document.

   A post-convening note that articulates a theory of digital market design stemming from architecture on the one end and mechanisms on the other will be written jointly with Mark Nottingham (IETF / Cloudflare).

2. **Foo Camp: "*Market Design and the AI Economy*"** (Lighthaven, Berkeley, 26–28 June 2026). Co-led with O'Reilly Media; 150–200 participants including Reid Hoffman (Greylock), Mitchell Baker (co-founder, Mozilla), Tino Cuéllar (Carnegie Endowment / Stanford; former Justice of the California Supreme Court), Kevin Kelly (Wired), Boris Cherny (Anthropic; Claude Code), and Isaiah Andrews (MIT). Leading philanthropies will be in attendance too.

3. **Mechanisms for Data Markets** (September 2026). Jointly with Microsoft Research and E. Glen Weyl, on the institutional design, valuation, and mechanism choices required to enable humans to actively participate in jointly produced human+AI data markets, thereby sharing in the co-created value.

Protocol-layer interventions: prototypes and research specifications

Few other policy-orientated organizations are as well positioned as the AI Disclosures Project to shape AI markets through bottom-up protocol and architectural interventions grounded in economic theories of market and systems design. To advance this agenda we are working on six flagship deliverables or implementations, each with external partners:

1. **Advancing a healthy skills marketplace.** Currently, it is difficult to tell what a skill covers (relative to some ideal set of capabilities) and if it is trustworthy. We aim to address this by focusing on two aspects of the skill marketplace: identity and skill coverage. For identity, we aim to work with people deeply embedded in the skill ecosystem, such as the MCP skills working group to think about linking skills to their creators for security and attribution. For skill coverage, we plan to widely release our own open source prototype of "skill cards", which will allow anyone to quickly compare skill content across multiple competing skills, relative to a reference set of capabilities.

2. **Federated protocols for agent-to-agent communication (e.g., ActivityPub and Nostr).** Existing agent-to-agent protocols lack a discovery mechanism and assume one-to-one

interactions; whereas ActivityPub (W3C) and Nostr (Jack Dorsey supported) already provide humans with stable identity, accumulated reputation, and native discovery across a decentralised network. Federated protocols allow agent identity to be made portable across hosts and prevents discovery from being captured by a single platform. Our prototype demonstrates this thesis on Nostr. When advancing it beyond the prototype stage, we will aim to collaborate with the W3C ActivityPub community, the Nostr ecosystem, the Linux Foundation's Agent2Agent (A2A) project, AGNTCY (Cisco's Internet of Agents initiative), and potentially Mark Nottingham at the IETF.

3. **Interoperable agent memory via ACP and MCP.** Building out an open source memory server to break down existing siloed and bundled memory systems (what we call the "memory walled garden"). Our prototype demonstrates cross-platform memory sharing through a standardized interoperable and modular memory system, implemented through MCP transport. Collaborations with the Goose team and the Agentic AI Foundation will be explored in order to bring full cross-platform chat history to coding agents.

4. **Using goose as a platform for implementation and prototypes**. Goose is a good fit for our work as it is a neutrally governed open-source MCP host. Goose has already played this role once: MCP UI found its footing in Goose before being adopted across OpenAI and Anthropic products as MCP Apps. We expect this sort of pattern to repeat itself as without a complete open-source agent host, new protocol features cannot be exercised in public, and closed commercial agents are ill-suited to serve as the venue for community experimentation.

5. **A "Neutral" AI Web Repository / Index**. Led by the AI Disclosures Project, with Creative Commons (partnership in discussion), the Knight-Georgetown Institute, and others, the proposal aims to disincentivize agentic web scraping by recreating a linked web of informational relevance for agents: creators register their work (or a pointer to it) to establish first-publication provenance, and have their relative informational uniqueness calculated. AI developers query the index or database that shows data uniqueness and priority instead of crawling the open web. Entries rank on relative distinctiveness (judged by their novelty, in terms of entropy / information gain), not popularity (PageRank). Content fingerprinting, a "Shazam for text," resolves copies and near-copies back to the original author. A non-commercial base layer (what exists, who made it first) is kept separate from a commercial layer (access and compensation), with patent-style registration friction – such as a small fee, strict formatting, or a self-reported novelty claim – to deter spam.

6. **Automated open protocol adoption scorecard.** Using capability advertisement features in MCP and ACP, we will measure the extent of open protocol adoption across AI platforms. This will go beyond assessing for binary MCP or ACP support to interrogating which protocol functions are and are not supported by a platform (such as sampling in MCP or session history in ACP). In addition to providing a public resource for developers, this tracking and evaluation scorecard will also let us create the kind of public visibility that pressures platforms to close those gaps, as well as advocating for specific protocol features where gaps may appear.

The above builds on our earlier impactful work on protocols and power, disclosures and attribution, and missing market mechanisms to include humans in AI markets, such as for training data consumption.

Operations and Donor Support. At our new fiscal sponsor, Code for Science & Society, our fringe rate is an economical 25% with overhead of 15%. We also benefit from renting highly subsidized office space at Collider in Financial District (New York), which supports an ecosystem of AI research and other startups. Visit our new website for more information.

4. Closing

The AI economy is being built right now. The mechanisms, protocols, and architectures being decided over the next 18–36 months will determine whether AI markets reproduce the winner-take-most dynamics of the platform era, or open up a genuine architecture of participation in which value flows back to the humans, institutions, and communities that produce, curate, and consume it.

Bellagio was an intensive working session that took seriously the possibility that the right small group of people, given proper framing and a few days together, can begin architecting healthy AI markets through key intervention points.

We are grateful to the Rockefeller Foundation for hosting Bellagio – and to our funders. The work is now ahead of us. We would welcome the chance to discuss how your foundation might join it. Highlights of Work Q4, 2025 – Present can be found in the Appendix below.

Appendix

Work Highlights from Q4, 2025 – Present

Convenings

Rockefeller-Sponsored Bellagio Center convening on **Missing Mechanisms for Human+AI Markets** (April 2026) — complete. Twenty participants across the IETF, MIT CSAIL, Yale, Microsoft Research, Block, Cloudflare, ex-Reddit, and Knight-Georgetown Institute. The output of Bellagio is the program described in this document.

Foo Camp (*forthcoming*): "*Market Design and the AI Economy*" (Lighthaven, Berkeley, 26–28 June 2026). Co-led with O'Reilly Media; 150–200 participants including Reid Hoffman (Greylock), Mitchell Baker (co-founder, Mozilla), Tino Cuéllar (Carnegie Endowment / Stanford; former Justice of the California Supreme Court), Kevin Kelly (Wired), Boris Cherny (Anthropic; Claude Code), Isaiah Andrews (MIT) and many others.

Mechanisms for Data Markets (September 2026, *day to be confirmed*). Jointly with Microsoft Research and E. Glen Weyl, on the institutional design, valuation, and mechanism choices required to enable substantive human participation in jointly produced human+AI data markets.

Presentations, panels, and chaired sessions

- **TRAILS Con 2026, "Transparency as a Tool for Accountability"** (George Washington University, Washington DC, 5 March 2026). Ilan designed and ran this session, chaired by Susan Ariel Aaronson (GW; TRAILS Co-PI, Governance Research Lead). TRAILS is the NSF/NIST-affiliated Institute for Trustworthy AI in Law & Society. Agenda.
- **Microsoft AI Economy Institute (AIEI) workshop on AI markets** (April 2026). Ilan chaired a session on AI markets at the in-person workshop convening of AIEI Senior Fellows at Microsoft. AIEI program.
- **Notion** (March 2026). Tim presented to the Notion CEO team on AI markets.

Press and media

- **Financial Times**, *"AI hallucinations haunt users more than job losses"* (22 March 2026). Ilan was cited extensively as the principal technical critical voice on Anthropic's 80,000-user Claude survey.

- **Financial Times, Swamp Notes** "*Humans Vs Bots*" (Rana Foroohar column, 26 January 2026). Tim wrote a piece on the AI economy as production without demand, applying Bostrom's paperclip thought experiment to corporate use of AI to eliminate workers while productivity gains flow to shareholders.

- **Partnership on AI**, *2026 Transparency Report on Foundation Model Impacts* (29 April 2026). The AI Disclosures Project is cited in PAI's progress report on post-deployment governance practices, which reviews more than 150 papers, articles, websites, and reports.

- **TechPolicy Press**, *"AI Isn't a Superintelligence. It's a Market in Need of Disclosure"* (Strauss & O'Reilly, 27 October 2025). Perspective piece arguing that AI risks belong in the SEC's existing disclosure machinery — 10-K, 10-Q, 8-K — rather than in a parallel regime built around science-fiction existential risk.

- **The Regulatory Review** (Penn Program on Regulation, University of Pennsylvania Carey Law School), *"Regulating AI Washing"* — Saturday Seminar (7 March 2026). Cites the Strauss, O'Reilly, Rosenblat, and Moure SSRC working paper *"Governing AI Through SEC Disclosure"* as one of six anchor scholarly sources in a meta-review of the field. The piece highlights our review of 7,800+ corporate filings finding that roughly two-thirds of corporate AI disclosures emphasise benefits while omitting significant risks, and its four-part reform proposal.

- **Big Think**, *"'Surfing the edge': Tim O'Reilly on how humans can thrive with AI"* (2 December 2025). Long-form feature interview with Tim on AI deployment choices, capital allocation, and the WTF Economy framework as it applies to the current AI cycle.

- **DesignRush Podcast**, Tim O'Reilly on why generosity is the ultimate growth strategy in AI and open source (26 September 2025). Tim argues that companies focused on short-term advantage by hoarding models risk being overtaken by more open, decentralised competitors – that "generosity and low barriers to entry create innovation" and that "systems that invite participation win." A compact restatement of the architecture-of-participation thesis applied to current AI market dynamics. Coverage.

Research and publications (published or forthcoming)

- "M&A as Capability Building: Patent Evidence from Big Tech" (Strauss, Yang), *Applied Economics Letters*, published **here**.

- "The Attribution Crisis in LLM Search Results" (Strauss, Yang, O'Reilly, Rosenblat, Moure), *Data & Policy* (Cambridge University Press), published **here**.

- "Real-World Gaps in AI Governance Research" (Strauss, O'Reilly, Rosenblat, Moure), *IEEE Access* — revision accepted. Draft **arXiv**.

- "Rich-Get-Richer? Platform Attention and Earnings Inequality Using Patreon Earnings Data" (Strauss, Yang, Mazzucato), forthcoming in *Industrial and Corporate Change* (accepted). Draft **arXiv:2509.26523**.
- "Beyond Public Access in LLM Pre-training Data" (Rosenblat, O'Reilly, Strauss), *AI & Ethics* — revision accepted. Draft **arXiv**.
- "Addictive or Helpful? AI Follow-up Questions and User Engagement" — draft forthcoming. Github.
- "Architectural Rents" — theoretical paper with Mariana Mazzucato on positional rents derived from a firm's position in a layered digital network, distinct from but contributing to gatekeeper and monopoly rents.

Asimov's Addendum — *ongoing thought leadership and commentary to 1.29k subscribers, as recommended by Helen Toner and authors of several other notable Substacks*

- The Third Artificial Intelligence – Tim O'Reilly, May 2026
- The Collaborative Exoskeleton of AI Science – Tim O'Reilly, May 2026
- Missing Mechanisms for the Agentic Economy – Tim O'Reilly, April 2026
- Don't Blame the Model (republished on *O'Reilly Radar*) – Sruly Rosenblat, March 2026
- Dangerously Skip Permissions – Sruly Rosenblat, February 2026
- Can Open Source Survive AI? – Sruly Rosenblat, February 2026
- What Does Anthropic's Latest Economic Index Data Really Tell Us? – Ilan Strauss, January 2026

Institutional

- Fiscal sponsorship transitioned from the Social Science Research Council (SSRC) to Code for Science & Society (March 2026), positioning the AI Disclosures Project for a multi-funder second round, broader technical / protocol-layer work, simplified backend systems, and with an official advisory committee now.
    - Lowered fringe rate from 33% to 25%
    - Switched to low-cost, highly subsidized office space at Collider in the Financial District, NY (26 Broadway, Floor 3, New York, NY 10004)
    - Rebranded with a new website