

The Subsidy That Isn't: Why Low-Cost AI Tools Hide High-Cost Exclusions

Weekly Analysis — <https://ainews.social>

The most successful trick the current generation of AI tools has pulled is convincing a few hundred million people that they are getting something for nothing. The chatbot is free. The coding assistant is free. The image generator is free, at least until you hit your daily cap. The meeting recap writes itself, and your employer pays for the seat anyway, so what's to complain about? This is the framing the vendors prefer, and it has done extraordinary work in normalizing a particular distribution of risk: the tool is positioned as a subsidy to the user, when in fact the user is a subsidy to the tool.

Watch how the freemium architecture is actually built. The free tier of every major large language model runs on smaller context windows, slower inference, older snapshots of the underlying model, and aggressive rate limits — a structure the vendors describe as "accessible AI" while quietly reserving the actually-functional product for paying customers. Microsoft's own product documentation makes this two-tier logic explicit in its corporate stack: Copilot's most powerful features in Business Central are gated behind specific license SKUs, with the "AI capabilities" page reading more like a procurement matrix than a feature announcement [2]. The "intelligent recap" of Teams meetings, marketed as one of the marquee productivity gains of the generative era, requires a Teams Premium or Copilot license to access the parts that anyone actually wants — speaker timelines, automatic action items, personalized summaries [19]. The free seat sees the room. The paid seat understands it.

None of this is hidden, exactly. It is simply not narrated, and the vendor-friendly trade press has been content to leave it un-narrated, because the alternative story — that the AI tool economy is rapidly stratifying into a two-tier system where the wealthy get private, fast, current models and the poor get slow, surveilled, stale ones — is harder to sell as innovation. The dominance of "tool/utility" framing in industry coverage (around a quarter of all AI tools discourse in the corpus this publication tracks) does the rhetorical work of presenting these objects as neutral implements, like hammers, that anyone can pick up. They are not hammers. They are subscription services with metered access, and the meter is built into the architecture.

[2] AI in Business Central

[19] Resumen inteligente de las llamadas y reuniones de Teams

What the Free Tier Actually Buys You

The most useful question to ask of any AI tool that calls itself free is: what is the degraded version of the product, and what is the relationship between the degradation and the price ladder above it? In the consumer chatbot market the gap is now substantial. Free users of the major assistants are routed to smaller or quantized models, given shorter context windows, throttled when traffic peaks, and cut off from the most recent model releases for weeks or months after the paid tier gets them. The vendor framing is that these are "fair use" limits. The structural reality is that the free product is designed to be visibly worse than the paid one in ways that are calibrated to drive conversion, not in ways calibrated to be the best free product the company could ship.

The developer tools market shows the pattern with unusual clarity because the products are technical enough that the gap can be measured. GitHub Copilot offers a free tier for students with what the documentation describes as "access to the same AI models that power the paid product" — but only if you can prove enrollment at a qualifying institution and only for as long as that enrollment lasts [1]. Outside that demographic window, the same student becomes a paying customer or loses the tool. Amazon's CodeWhisperer pitched itself as an individual-tier alternative with an aggressive free offering, but its actual user guide reveals an architecture in which the meaningful security and compliance features are reserved for the Professional tier [5]. Google's Gemini for Developers tutorials walk new users through the API with generous quota allotments, but those allotments are structured to terminate, not to scale [13]. Microsoft's own AI-assisted development push inside Visual Studio bundles the most useful features into the licensed product and reserves the demonstration-quality completions for free [6].

This is not a complaint about commerce. Software companies are allowed to charge for software. The complaint is about the rhetorical move that lets these products be described as a democratization of AI — as if the tier that low-income or low-status users can actually reach is the same tool, only smaller. It is not the same tool. It is the trial-sized box of detergent placed at eye level so that the full box across the aisle looks reasonable by comparison. The Renaissance Society subscriber considering whether to sign her grandchildren up for "free" Copilot access should know that the free version is a marketing instrument before it is a tool, and that the conversion funnel is the actual product the vendor is shipping.

[1] Access GitHub Copilot for free as a student

[5] Amazon CodeWhisperer Documentation

[13] Gemini for Developers | Google Codelabs

[6] Assistance par intelligence artificielle pour les développeurs dans l
...

The Data Commons Only One Party Can Sell

The second cost of the free tier is data, and here the vendor narration becomes especially slippery. The user-facing message is that free-tier inputs “may be used to improve our services.” The corporate reality, where it has been documented at all, is that those inputs constitute the training and fine-tuning substrate for the next generation of paid products. The free user is paying in a currency — text, code, voice, behavior, correction signals — that the company can monetize and the user cannot.

GitHub’s recent move to use customer code to train Microsoft’s models, unless the customer affirmatively opts out, is the cleanest recent example. The Spanish technical press captured the move with appropriate bluntness: GitHub will use your code to train Microsoft’s AI if you do not explicitly stop it [14]. The default is extraction. The opt-out is a buried setting. The user who never finds it is contributing labor to a corpus they cannot inspect, audit, or be paid for. Shoshana Zuboff’s account of behavioral surplus — the data exhaust that capitalist firms convert into prediction products — was written about an earlier generation of platforms, but the architecture has migrated cleanly into the generative tools layer [7]. What is novel is the speed: the round trip from user prompt to model improvement to upsell pitch can now close inside a single quarter.

Kate Crawford’s mapping of the AI supply chain is useful here too, because it refuses the politeness of the term “data” [7]. The text a free-tier user types into a chatbot is labor — the labor of articulating a problem, the labor of correcting the model’s mistakes, the labor of demonstrating what a good answer looks like. Calling it data launders the relationship. The corpus that emerges is not a natural resource that the company has merely collected; it is a commons that the company has unilaterally enclosed. And the enclosure is one-directional: the user cannot extract value from the corpus their inputs helped build, cannot license it to a competitor, cannot even reliably delete their contribution. The free tier is the mechanism by which this enclosure proceeds at scale, because pricing the service at zero is the only way to recruit the volume of labor required.

The privacy stakes compound when the tools cross into institutional settings where the users are minors or are otherwise unable to consent in any meaningful sense. Coverage of AI use in schools has documented a sharp rise in data breaches, deepfake incidents, and downstream harms as the tools have proliferated without commensurate data-governance infrastructure [4]. The French government’s back-to-school dossier on AI in 2025 makes the regulatory gap unusu-

[14] GitHub va a usar tu código para entrenar la IA de Microsoft ... si no ...

[7] The Age of Surveillance Capitalism

[7] The Atlas of AI

[4] AI Use in Schools Soars as Data Breaches, Bullying and Deepfakes Rise

ally explicit, noting that the consumer tools children are using daily were not designed for, and have not been audited against, the data-protection standards that apply to educational records [8]. The free tier, in other words, is not merely surveillance; it is surveillance that has slipped past the institutional review that would normally apply to a comparable extraction of personal information.

Microsoft’s research division has begun publishing work on privacy-preserving alternatives — Vega, a recent zero-knowledge-proof framework for digital identity in AI contexts, is presented as a way to verify properties of users without exposing the underlying data [21]. It is worth noticing where this kind of engineering lives in the corporate org chart. It is in research, not in product. The same company that publishes papers on cryptographic privacy guarantees ships consumer products whose default settings Hoover user data into training corpora. The asymmetry is not an accident; it is the business model.

The Open-Source Escape Hatch That Isn’t

The standard response to the freemium critique is that the open-weight alternatives exist. Llama, Mistral, Qwen, DeepSeek — anyone who objects to the proprietary tools can run their own. This response collapses on contact with the actual hardware requirements and the actual technical learning curve.

A useful adopter test: can you, the reader of this essay, get a recent Llama-class model running locally on a laptop you already own, with a context window comparable to the paid tier of the proprietary product, at usable speed? For the vast majority of educated non-specialists, the honest answer is no. The hardware investment to run a competitive open-weight model at home now sits in the multi-thousand-dollar range for a usable GPU setup, and the software stack — quantization, inference engines, prompt templates, retrieval augmentation — assumes a baseline of systems-level fluency that the free-tier ChatGPT user has not been required to develop. The open-source escape hatch is real, but it is a hatch that opens upward, toward people with technical capital, not downward, toward people excluded by financial capital.

This matters for the equity argument because it means the “alternative” to the surveilling, throttled, monetizing free tier is not actually available to the demographic that the free tier is supposedly serving. Meredith Broussard’s category of technochauvinism — the belief that technical solutions are always preferable, even when they are not technically accessible to the people whose problems they are claimed to

[8] Back to school in 2025: teaching and learning in the age of AI and the ...

[21] Vega: Zero-knowledge proofs for digital identity in the age ...

solve — fits the open-source escape hatch with depressing precision [7]. The user told to “just run Llama locally” is being told, in effect, to become a different person before her tool problem can be solved.

[7] Artificial Unintelligence

The funding picture compounds the access gap. Proprietary AI labs operate on tens of billions of dollars of capital. The open-source ecosystem operates on a small fraction of that, distributed across academic groups, hobbyist contributors, and a handful of corporate releases whose business case is strategic rather than philanthropic. Meta releases Llama because it weakens its competitors’ moat, not because it has discovered an interest in equity. When the strategic logic changes — and it will — the release cadence will change with it. Ruha Benjamin’s analysis of how technical infrastructures encode the priorities of the actors who fund them is directly relevant: the open-weight commons exists at the pleasure of corporate actors whose interests are tactical, and its continued availability is not a property of the technology but a property of the current market structure [7].

[7] Race After Technology

The interoperability story is similarly thin. Proprietary tools have made it progressively harder to migrate work between platforms — prompt libraries, fine-tunes, custom GPTs, and stored conversations all live in formats designed to retain users, not to liberate them. The user who has invested several hundred hours in calibrating her workflow inside one vendor’s tool faces a switching cost that the next vendor’s marketing copy is determined to obscure. This is the lock-in that the free tier is designed to produce: the price of egress, once you have grown dependent, exceeds the price of staying.

Performance Drift and the Tools That Quietly Get Worse

A second category of hidden cost, less discussed but well-documented in the technical press, is performance drift. The model you signed up for is not the model you are using six months later. Vendors retrain, fine-tune, add safety layers, change default temperature settings, and route requests through different infrastructure as load demands shift. The user experience can degrade — sometimes dramatically — without any change in price, contract terms, or service description.

The French science magazine *Epsilon* recently published a careful examination of the gap between what the major generative models claim to do and what they actually do, finding consistent overpromising on reliability, factuality, and reasoning [15]. The reporting found that benchmarks the vendors cite tend to be selected from periods when the model performed at its peak, with the inevitable post-deployment drift quietly absorbed by users. The free tier, where

[15] IA : La faille | Epsilon

the user has the weakest contractual position, is where this drift is most easily imposed. There is no SLA. There is no compensation. The product simply works less well one Tuesday than it did the previous Tuesday, and the user is invited to upgrade.

The detector ecosystem that has grown up around these tools offers a particularly grim window into performance claims. Princeton's recent work on AI-detection accuracy found roughly three times the false-positive rate the vendors had advertised, with downstream effects on the people — disproportionately non-native English speakers — flagged as cheats [9]. The follow-up reporting on causes and impacts of false positives in the detection layer makes clear that the failure mode is not a bug to be fixed in the next release; it is a structural property of the way the detectors were built [11]. The tools claim a precision they cannot demonstrate, and the cost of the false claim is borne by the wrong end of the power gradient.

Even in the marquee research narratives that vendors lean on to justify the freemium expansion, the evidence is messier than the press releases suggest. A widely circulated Nature study on AI tutoring outperforming in-class active learning is real and worth reading, but its design — a randomized controlled trial in a specific Harvard physics course with researchers building bespoke tooling — bears very little resemblance to the consumer-grade chatbot a free-tier user is actually being sold [3]. The result was achieved with a carefully engineered system. The marketing implication is that any free chatbot will do. The gap between the two claims is precisely the gap the vendor wants the reader not to notice.

The Consolidation Beneath the Abundance

The third structural cost of the current tools market is the consolidation that the freemium tier funds. The reason vendors can afford to give the product away to hundreds of millions of users is that a small number of them subscribe, that the data flywheel improves the paid product, and that the capital markets are willing to subsidize the loss in pursuit of monopoly position. The result, three years in, is a tools market dominated by a handful of firms whose pricing power is rapidly extending into every adjacent layer of the stack.

Microsoft's own recent strategic anxiety is worth attending to. Coverage of Copilot's stumbles in the broader AI race has portrayed Satya Nadella's company as having "lost its way" precisely because the consolidation strategy did not deliver the user-base dominance the company expected; Google and Anthropic are now eating into share

[9] Détecteurs IA : Princeton Révèle
3x Plus de Faux Positifs

[11] Faux positifs détecteurs IA :
causes, impacts et solutions

[3] AI tutoring outperforms in-class
active learning: an RCT ... - Nature

that Microsoft assumed it had locked up by partnering with OpenAI [17]. This is not a story of competition restoring user welfare. It is a story of three or four firms competing for the position from which they will extract rent from the entire downstream tools market. The end state, in any of the plausible scenarios, is a tools layer controlled by a small oligopoly with no meaningful exit option for the dependent user.

Adobe’s repositioning of Firefly as “your all-in-one creative AI studio” is a textbook example of how this consolidation looks at the application layer [10]. The independent generative-image tools that proliferated in 2023 are being absorbed, undercut, or rendered irrelevant as Adobe folds equivalent functionality into a subscription bundle that the creative-professional class is already locked into. The Stable Diffusion copyright cases that introduced the legal earthquake risk to open-source generative imaging have done significant work to clear the field for incumbents who can absorb litigation cost [20]. The result is that the creative-tools market is moving toward a configuration in which a few large vendors decide what generative tools look like and on what terms.

Virginia Eubanks’s work on how automated systems entrench existing inequalities is directly applicable here, even though her cases involved welfare and policing rather than generative chatbots [7]. The mechanism is the same: a small number of vendors get to set the defaults, the defaults encode the priorities of the vendor, and the populations most dependent on the system have the least leverage to contest those defaults. The freemium chatbot is the welfare-eligibility algorithm of the generative era — universally available, structurally tilted, and rhetorically presented as a neutral service.

What the Tools Actually Do, Versus What’s Claimed

The analytical question this essay is built around — what do AI tools actually do, versus what’s claimed — has a sharper answer than the trade press is willing to commit to. The tools do narrower, less reliable, more contingent things than their marketing implies; they do those things on terms that extract value from users who cannot afford to pay; and they are converging on a market structure in which the user’s leverage will continue to decline.

The skeptical-versus-PROAI *stance distribution in tools coverage tells its own story*. The PROAI corpus is dominated by vendor product announcements, partnership press releases, and explainer pieces that treat tool capabilities as straightforwardly described by the vendors who built them. The skeptical corpus is smaller, often relegated

[17] Microsoft lost its way in the AI race. Can Copilot get it back on course?

[10] Explore the new Adobe Firefly, your all-in-one creative AI studio

[20] Stable Diffusion copyright lawsuits could be a legal earthquake for AI

[7] Automating Inequality

to opinion pages, and tends to focus on downstream harms rather than upstream design choices. The American Academy of Arts and Sciences' framing of AI as simultaneously a social-science question, object, and tool is one of the few high-status interventions that takes seriously the idea that these are not neutral instruments [12]. Most of the coverage the careful adopter encounters has not made that move. It treats the tool as a hammer.

The skepticism that does break through tends to come from the institutional periphery rather than the center. Recent NPR coverage of an extensive report finding the risks of AI in schools outweigh the benefits stands out partly because the public-radio register can accommodate an analytic verdict that the tech press will not deliver [18]. CNN's reporting on the Alpha School experiment — an entire institutional concept built around the premise that AI can replace teachers — captures the bet structure underneath the rhetoric: a risky wager being placed with other people's children as the stake [16]. The tools are not being neutrally evaluated. They are being deployed at scale, on populations who cannot opt out, with the evaluation deferred to whoever survives the deployment.

The Questions a Careful Adopter Should Actually Ask

If the freemium architecture is a subsidy that runs in the wrong direction, the prudent reader's response is not abstinence but interrogation. The questions worth asking of any AI tool that lands in your hands, regardless of the price tag, are concrete enough to be useful.

What is the gap between the free tier and the paid tier, and where exactly does that gap fall — in latency, in model size, in context window, in update cadence, in feature availability? If the vendor cannot tell you, that is itself the answer.

What happens to your inputs? Are they used to train the next model? Can you opt out? Is the opt-out the default or buried in a settings menu three layers deep? If it is buried, the vendor has told you what they think of your consent. The GitHub case is the model: extraction is the default, opt-out is the exception, and the burden of awareness is dumped onto the user [14].

What is the switching cost? If you decide in eighteen months that this tool was a mistake, can you take your prompt library, your fine-tunes, your stored conversations, and your workflow somewhere else? Or has the platform been deliberately engineered to make egress expensive?

[12] Field Theory: AI as Social Science Question, Object & Tool

[18] Report: The risks of AI in schools outweigh the benefits : NPR

[16] Is AI schooling the future of education — or a risky bet? - CNN

[14] GitHub va a usar tu código para entrenar la IA de Microsoft ... si no ...

What is the vendor’s actual financial structure? Is the free tier funded by paid subscriptions that are themselves profitable, or by venture capital that will run out, or by a strategic loss-leader play that will end the moment the strategic logic shifts? If the answer is the second or third, the tool you depend on is on a clock the vendor controls.

What is the failure mode? When the tool is wrong, who bears the cost? The detector case is the cleanest example: the tool’s false positive is the user’s academic discipline hearing [11]. The vendor has externalized the risk. The careful adopter notices.

[11] Faux positifs détecteurs IA : causes, impacts et solutions

And what is the institutional alternative if the vendor exits the market or doubles the price? If there is no alternative — if you have been routed onto a single proprietary stack with no exit — you are not a customer in any meaningful sense. You are a hostage with a recurring credit-card charge.

These are not paranoid questions. They are the standard questions one would ask of any utility on whose continued operation one’s work depends. The reason they sound paranoid in the AI tools context is that the vendors have worked hard to make the tools feel like consumer products rather than utilities, and the trade press has been content to validate that framing. Stripping the framing is the first thing the careful adopter has to do.

The Subsidy Runs the Other Way

The titular claim of this essay is that the low-cost AI tool is not a subsidy to its user. The argument has been that the user is the subsidy — paying in data, attention, switching cost, and lock-in for a product that has been engineered to make those payments invisible. The architecture is not accidental. It is the business model. And the people for whom the free tier is most consequential — students, workers, freelancers, the credentialed but cash-poor knowledge class that has been told these tools are democratizing — are the people whose subsidy is largest and whose leverage is smallest.

There are redesigns that would interrupt this dynamic. Privacy-preserving inference, local-first models, interoperability standards, data-portability requirements, genuine open-weight ecosystems with real funding — each of these has technical proofs of concept and policy advocates. None of them is incompatible with a healthy tools market. All of them are incompatible with the current business model, which is why the vendors are not building them and the regulators have not yet required them. The Vega work at Microsoft Research is

the proof that the engineering exists; the absence of Vega-like features in the consumer products is the proof that the engineering is not the constraint [21]. The constraint is the business model.

The honest version of the freemium pitch would read something like this: this tool is free because we have determined that your data and attention are worth more to us than what we could charge you in cash; the version of the product we will sell you costs money because that version is the one we have made functional; the version we will sell to your employer costs more because that version is the one we have made auditable; and the architecture is designed to make leaving us expensive enough that you probably won't. No vendor will write that pitch. The careful adopter has to write it for herself, and then decide, with the pitch on the table, whether the trade is one she wants to make.

What she should not do is accept the framing that the tool is a gift. The Renaissance Society subscriber who has read this far will recognize the structure: it is the same structure that magazines have used to sell subscriptions, that grocery stores have used to sell loyalty cards, that platforms have used to sell themselves into every previous decade's communications infrastructure. The novelty is the speed of the lock-in and the depth of the data extraction. The novelty is not the trick. The trick is old. The pricing tier called "free" has always been the most expensive one. The current generation of AI tools has only made the price harder to read.

References

1. Access GitHub Copilot for free as a student
2. AI in Business Central
3. AI tutoring outperforms in-class active learning: an RCT ... - Nature
4. AI Use in Schools Soars as Data Breaches, Bullying and Deepfakes Rise
5. Amazon CodeWhisperer Documentation
6. Assistance par intelligence artificielle pour les développeurs dans l ...
7. Automating Inequality
8. Back to school in 2025: teaching and learning in the age of AI and the ...

[21] Vega: Zero-knowledge proofs for digital identity in the age ...

9. Détecteurs IA : Princeton Révèle 3x Plus de Faux Positifs
10. Explore the new Adobe Firefly, your all-in-one creative AI studio
11. Faux positifs détecteurs IA : causes, impacts et solutions
12. Field Theory: AI as Social Science Question, Object & Tool
13. Gemini for Developers | Google Codelabs
14. GitHub va a usar tu código para entrenar la IA de Microsoft ... si no ...
15. IA : La faille | Epsilon
16. Is AI schooling the future of education — or a risky bet? - CNN
17. Microsoft lost its way in the AI race. Can Copilot get it back on course?
18. Report: The risks of AI in schools outweigh the benefits : NPR
19. Resumen inteligente de las llamadas y reuniones de Teams
20. Stable Diffusion copyright lawsuits could be a legal earthquake for AI
21. Vega: Zero-knowledge proofs for digital identity in the age ...