The 18-Watt Expert: GPT-5's Launch and the Materiality It Erases

Weekly Analysis — https://ainews.social

When OpenAI told the world that using GPT-5 was "like talking to a legitimate PhD-level expert," the sentence did more work than any benchmark chart in the launch deck. It converted a statistical artifact—a very large function that predicts the next token—into a person with a credential, a discipline, and the authority those things imply. You do not audit a PhD. You do not ask where a PhD's electricity comes from, or whether the PhD is quietly forwarding your questions to a data center you have never heard of. You defer. That is what the metaphor is for. The professor's gown is not a description of the machine; it is a request that you stop looking at the machine.

This is the oldest trick in the field, and it has a name. The people who build these systems have always understood that a chatbot passes as human not by being intelligent but by managing your expectations—in one now-canonical case, an early bot posed as an eleven-year-old Ukrainian boy with poor English, a costume designed so that its non sequiturs would read as charm rather than failure [22]. GPT-5's costume is grander—a doctorate rather than a child—but the mechanism is identical: supply a frame that explains away the seams before you notice them. Kate Crawford calls the deeper version of this move enchanted determinism, the practice of directing public attention "to the innovative nature of the method rather than on what is primary: the purpose of the thing itself" [22]. Enchantment, she writes, "obscures power and closes off informed public discussion, critical scrutiny, or outright rejection."

[22] You Look Like a Thing and I Love You

[22] The Atlas of AI - Power, Politics, and the Planetary Costs

So let us decline the enchantment and ask the primary questions. What does this tool actually do? What does it do *to* you? Where does the bill go, and who pays it? Because the striking thing about the launch week was not the model. It was how much of the material reality—the watts, the data flows, the security surface, the ownership—kept breaking through a narrative engineered to keep it invisible. The tool is a power plant wearing a professor's gown, and the gown is the product.

The Metaphor Does the Hiding

Start with the word ”expert,” because everything else is downstream of it. An expert is accountable to a body of knowledge and to peers who can sanction error; an expert can say ”I don’t know”; an expert has a history you can check. A language model has none of this. It has a training distribution and a temperature setting. When it produces a confident falsehood—and it will—there is no discipline to which it answers and no mechanism by which it learns it was wrong in your particular case. Meredith Broussard’s term for the underlying confusion is technochauvinism: the reflex that a computational solution is inherently superior, a reflex that survives precisely because we keep asking machines to ”understand” things that they are, in fact, only processing [22]. The ”PhD-level expert” line is technochauvinism compressed into marketing copy.

[22] Artificial Unintelligence - How Computers Misunderstand

Watch how the frame travels. Once a model is an expert, its outputs are consultations; once they are consultations, the appropriate posture is trust; and once trust is the posture, the questions a careful adopter should be asking—about provenance, cost, and failure modes—start to feel rude, even Luddite. This is not accidental. The metaphor performs the concealment. It is worth noticing how differently the same industry speaks when it is billing you rather than selling you. In the pricing documentation, there are no experts and no doctorates; there are only tokens, metered to the fraction of a cent, input distinguished from output, cached distinguished from fresh [20]. The enterprise tarifario is a masterpiece of granularity, a document that knows exactly how much of the machine you consumed [21]. The vendor can count your consumption to a resolution finer than a grain of rice. Hold that fact; we will need it in a moment, when we ask why the one number the vendor cannot seem to produce is the one measured in watts.

[20] Tabla de tarifas de ChatGPT (precios Enterprise basados en tokens)

[21] Tarifario de ChatGPT (Business, Enterprise/Edu) - OpenAI Help Center

The dominance of this ”tool” framing is not a vibe; it is a measurable feature of how the technology is discussed. Across the week’s coverage, the utility register—AI as tool, as assistant, as productivity multiplier—was the single most common way of talking about these systems, crowding out the registers that would foreground cost, risk, or power. That is what a successful mystification looks like from the outside: not the absence of skeptical voices but their reduction to a minority, a garnish on a plate that is mostly hype. The HAI AI Index’s own survey of the discourse captures the flavor of the ecstatic majority—developers announcing they can ”now build *complete* apps with nothing more than my voice. The future is here!”—the kind of frictionless testimonial that functions as free marketing and sets the

emotional baseline against which any caution reads as sourness [22].

[22] HAI_AI-Index-Report-2024

The point is not that GPT-5 is useless. It is that ”expert” is a claim about kind, not degree, and the evidence supports only a claim about degree. A better model is a better next-token predictor. That is genuinely valuable and genuinely limited, and the whole rhetorical apparatus of the launch exists to prevent you from holding both halves of that sentence at once.

The Bill Nobody Meters

Here is the question the professor’s gown is designed to make you forget: what does one answer cost the planet? Not in dollars—OpenAI will tell you that to the penny—but in energy. And here the vendor, so fluent in the arithmetic of tokens, goes conspicuously silent. OpenAI declined to disclose GPT-5’s energy consumption at launch. The company that can meter your usage to the fraction of a cent cannot, or will not, tell you how many watt-hours it burned to do so.

Into that silence stepped independent estimation. Researchers working out of Rhode Island put the figure for a single medium-length GPT-5 response at roughly eighteen watt-hours—enough to run an incandescent bulb for the better part of a hour, spent on one reply. Take that number, which the vendor has neither confirmed nor rebutted, and multiply it by the hundreds of millions of queries these systems field daily, and the ”expert” resolves into what it materially is: an industrial process. Crawford’s whole project in *The Atlas of AI* is to force this resolution, to drag the conversation away from the enchanted surface and toward ”the political economies of their construction” and ”the wider planetary consequences” that the enchantment is built to obscure [22]. The sector’s aggregate carbon footprint has grown into the range of a major industry—comparable, by several accounts, to commercial aviation—and it is climbing with every model that gets described as smarter, which in practice means larger, which in practice means hungrier.

[22] The Atlas of AI - Power, Politics, and the Planetary Costs

Notice the structure of the deal. The gains from GPT-5 accrue to the vendor and, secondarily, to the individual user who gets a faster draft. The costs—the electricity, the water for cooling, the strain on grids—are socialized, spread across everyone who shares the atmosphere and the power infrastructure, most of whom never typed a prompt. This is the classic shape of an externality, and the ”expert” metaphor is what keeps the externality off the invoice. A power plant that called itself a power plant would face an environmental review. A power plant that calls itself a PhD faces a product launch.

The compute-maximalist logic driving all this deserves to be named plainly: the industry's answer to almost every question is *more*. More parameters, more training data, more inference, more agentic autonomy running more queries in the background. The pricing structures reward exactly this—usage-based billing means the vendor's revenue rises with your consumption, which means the business model is structurally opposed to efficiency in any sense that would reduce the meter [14]. And because the energy line item never appears, there is no counter-pressure, no signal that would make a user or an enterprise ask whether the eighteen-watt-hour answer was worth eighteen watt-hours. You cannot economize on a cost you are not shown. The unmetered bill is not an oversight; it is the design.

[14] Modelos y precios para GitHub Copilot

There is a grim irony in how the security economics get sold alongside all this. Vendors reciting the return-on-investment case for AI-augmented defense lean on figures like IBM's estimate that organizations using security AI and automation extensively save on the order of $1.76 million per breach relative to those that don't. It is a real number and a real argument. But it is an argument conducted entirely in the currency of the balance sheet—breach costs avoided, analyst hours saved—and it never once touches the electricity line, the water line, or the socialized carbon line. The savings are private and legible; the costs are public and, by careful arrangement, illegible.

The Silent Substrate

If the energy question shows what the launch narrative erases about the physical world, the coding-agent question shows what it erases about your data. This is where the "expert" metaphor becomes most quietly dangerous, because a consultant you trust is a consultant you hand your files to.

Read the documentation for the major coding assistants as a genre and a pattern emerges. GitHub Copilot's docs are extensive on capability—completions, chat, agents that can traverse a repository—and organized overwhelmingly around what the tool can do for your velocity [5]. Gemini Code Assist is presented as an overview of features and integrations, a productivity surface bolted into your editor [9]. Amazon's assistant, having consumed the tool once branded CodeWhisperer into the larger Amazon Q Developer, is likewise documented as a set of powers [1]. Even the story of that consolidation—CodeWhisperer folded into a bigger platform—reads as a feature announcement rather than what it also is, a narrowing of the market into fewer, larger hands [4]. Microsoft's guidance for AI-assisted de-

[5] Documentación de GitHub Copilot

[9] Gemini Code Assist overview | Google for Developers

[1] Amazon CodeWhisperer Documentation

[4] CodeWhisperer is becoming a part of Amazon Q Developer

velopment in Visual Studio follows the same contour: here is what the assistant can build with you [3].

What these guides share is a structuring absence. The prominent narrative is capability; the buried or missing narrative is *where your code goes when the agent reads it*. When a coding agent "understands" your codebase, it is transmitting that codebase—your proprietary logic, your embedded secrets, your customers' schemas—to a remote model for inference. What is retained, for how long, under whose jurisdiction, and whether it can surface in a future completion for another user: these are the questions that determine whether the tool is safe to point at anything you don't want to publish. And they are precisely the questions the workflow documentation treats as edge cases, if it treats them at all. The consumer-tier Gemini Code Assist documentation is more concerned with the deprecation and migration path of individual accounts than with a frank accounting of the data relationship [8], and the marketplace listing sells convenience above all [7].

The same pattern governs the office-productivity tier, where the mystification is smoothed into the language of "adoption." Microsoft's Copilot materials are a case study in the infrastructure narrative that omits the infrastructure. There is an adoption guide, an enablement resource set, a formal adoption report template, a rollout runbook with minimum requirements—an entire bureaucracy of onboarding [11]. There is a dashboard to measure how thoroughly your workforce has taken up the tool [12]. There is a rollout guide keyed to organizational requirements [18]. And there is the reference architecture for building your own agents on Power Platform and Copilot Studio, a document that renders the plumbing as clean boxes and arrows [16]. What all this adoption apparatus measures is uptake—seats filled, features touched—never exposure. You are told how to get everyone using the assistant and given a template to prove you succeeded; you are not given a template to measure how much of your organization's data has crossed into the model's context window and what happens to it there. "What is Microsoft 365 Copilot?" the documentation asks, and answers with capability, integration, and the promise of grounded, tenant-aware intelligence [26]. The literacy that would let you evaluate the trade is simply not the literacy the guides are designed to build.

Even the most banal ambiguities go unresolved in the enthusiasm. Ask the vendor's own support forum whether images a model generates for you are yours to sell commercially, and you find genuine uncertainty about ownership and rights—a foundational business question treated as an afterthought [2]. If the industry has not settled who owns the outputs, the "expert" framing is doing even more concealing

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

[8] Gemini Code Assist consumer accounts | Google for Developers

[7] Gemini Code Assist - Visual Studio Marketplace

[11] Microsoft 365 Copilot adoption guide and overview for IT admins

[12] Microsoft 365 Copilot adoption report | Microsoft Learn

[18] Rollout Microsoft 365 Copilot to your organization

[16] Overview of Power Platform and Copilot Studio reference ...

[26] ¿Qué es Microsoft 365 Copilot? | Microsoft Learn

[2] Are the images generated by openai dalle commercially available and who ...

than we thought.

Prompt Injection and the Literacy That Matters

Now put the two erasures together—an agent that ingests your data and a model marketed as a trustworthy expert—and you arrive at the security reality the launch narrative is least equipped to discuss. Because the thing about an ”expert” who will read any document you show it is that anyone who can get a document in front of it can issue it instructions.

This is prompt injection, and it is not a fringe concern. It sits at the top of the OWASP list of security risks for large-language-model applications—the number-one threat, ahead of everything else, in the reference taxonomy security professionals actually use. The mechanics are almost embarrassingly simple: because the model cannot reliably distinguish the developer's instructions from content encountered in the wild, an attacker can bury commands inside a web page, an email, a calendar invite, or a code comment, and the ”expert” will dutifully obey the poison as readily as the prompt [27]. Point a Copilot agent at your inbox and an adversary who can send you mail can, in principle, address your assistant. The same architectural openness that makes these tools feel magically capable—they'll read anything, they'll act on anything—is the vulnerability.

[27] ¿Qué es un Ataque de Inyección de Prompt en IA? La Amenaza Oculta que ...

The threat runs deeper than runtime manipulation, into the training substrate itself. Data poisoning—the deliberate seeding of a model's training corpus with corrupted material—can implant behaviors and backdoors that lie dormant until triggered, an attack surface that is invisible precisely because it operates upstream of anything a user can see [6]. You cannot inspect the training data of GPT-5; it is a black box, and its opacity is sold to you as proprietary advantage. But that same opacity means you are trusting not only the vendor's competence but the integrity of every source the vendor scraped. The ”expert” may have been quietly miseducated, and neither you nor the vendor can fully audit the curriculum.

[6] Empoisonnement de données : la menace invisible qui cible les grands ...

The industry's own response to all this is instructive, because it confirms that the problem is real while quietly conceding it is unsolved. Cloud platforms now ship risk-and-safety evaluators—tooling to score generative outputs for various harms—as a standard part of the deployment stack [17]. There are playbooks for operationalizing security and safety evaluations across the model lifecycle [19]. There are governance frameworks for securing AI agents across an organization, an implicit admission that an autonomous agent is a thing that must

[17] Risk and Safety Evaluators for Generative AI - Microsoft Foundry
[19] Safeguarding LLM security & safety evaluations | Microsoft Learn

be *contained* [10]. These are serious engineering responses. But notice that they are all mitigations layered *around* an artifact that cannot itself be made trustworthy—guardrails bolted to a machine whose core failure mode is that it does what it is told by whoever manages to tell it. You do not need to govern and evaluate and contain a genuine expert. The security apparatus is the vendor's own confession that the metaphor is false.

[10] Gérer et sécuriser les agents IA au sein de l'organisation - Cloud ...

This is why the version of "AI literacy" now being marketed—and, increasingly, mandated—is close to worthless where it matters. In its dominant institutional form, *alfabetización en IA* is satisfied by checkbox documentation: a training module completed, an acknowledgment signed, a policy filed. That kind of literacy teaches people to use the tool confidently, which is exactly the disposition the mystification wants to cultivate. The literacy that actually protects you is nearly its opposite: the trained suspicion that recognizes when the model is being manipulated, that treats every agentic action on untrusted input as a potential attack, that asks where the data went before asking how fast the draft came back. That is a skeptical literacy, and no vendor's compliance module is going to build it, because it points directly at the vendor.

Who the System Actually Serves

Crawford's sharpest question is the one the enchantment is built to suppress: "Whom do these systems serve?" [22]. Strip the professor's gown and the answer is not mysterious. The energy costs are socialized; the data flows enrich the platform; the security surface is your problem to govern; and the ownership questions resolve, when they resolve at all, in the vendor's favor. Every arrow of value points one way.

[22] The Atlas of AI - Power, Politics, and the Planetary Costs

The consolidation is visible in the plumbing. When enterprises want to deploy a competitor's model—say, running Anthropic's Claude inside Microsoft's Azure—they hit quota walls and deployment failures, friction that quietly channels usage back toward the platform's preferred stack [24]. Access to the best capabilities is gated and tiered, expanded and rescinded at the provider's discretion, as the shifting entitlements around Gemini's various tiers demonstrate [13]. The pattern of lock-in is not incidental to the business; it is the business. Every "expert" metaphor hardens it, because dependence on a trusted advisor is stickier than dependence on a mere tool, and the whole adoption apparatus—the enablement guides, the rollout runbooks—is engineered to deepen that dependence before anyone has priced the

[24] Unable to Deploy Anthropic Claude Models in Azure AI - Quota ...

[13] Mises à niveau et limites des applications Gemini pour les abonnés ...

exit.

The people at the sharp end of this economy see the extraction clearly, because it is happening to them. Voice actors fighting the encroachment of AI dubbing in Hollywood are not confused about whether these systems are neutral tools; they are watching their recorded performances become the training feedstock for the machines that will replace them, value flowing from the labor that created it to the platform that captured it [25]. And the surveillance dimension is now explicit enough that it has become a product-differentiation axis: an AI notetaker can advertise, as its distinguishing virtue, that it *won't* sell the record of your meetings to your boss—which tells you exactly what the default expectation has become for the ones that don't make the promise [23]. The Zuboffian logic—that your behavioral exhaust is the raw material and you are not the customer—is now the water these tools swim in.

[25] Why voice actors are fighting Hollywood AI - Rest of World

[23] This AI notetaker won't sell surveillance to your boss

To be fair to the evidence, not every skeptical claim survives scrutiny, and a committed-in-the-reader's-interest posture has to say so. When a widely circulated study was taken to prove that ChatGPT was "making users dumber," a closer reading showed the finding was far more limited than the headline—the research did not support the sweeping cognitive-decline narrative that spread on its back [15]. This matters, and not only as a fairness gesture. It matters because the skeptical case does not need the exaggerations. The genuine problems—unmetered energy, silent data transmission, top-ranked injection vulnerability, ownership ambiguity, deepening lock-in—are all documented, all material, and all sufficient. Reaching for the shaky "makes you dumber" claim actually weakens the critique by tying it to a finding that won't hold, and the mystifiers are only too happy to debunk the weak claim and imply the strong ones fall with it. The discipline of the careful adopter cuts both ways: distrust the vendor's enchantment *and* distrust the counter-hype that overreaches.

[15] No, ChatGPT Isn't Making You Dumber—What the Study Really Reveals

What the Careful Adopter Should Actually Know

So return to the launch, and to the sentence that started it. "Like talking to a legitimate PhD-level expert" is not a lie exactly; it is a redirection, a request to evaluate the artifact by its most flattering surface and to leave everything underneath in the dark. The essay's whole argument is that the underneath is where the decisions live.

A careful adopter, offered GPT-5 or any of its rivals, should hold four questions in mind that no launch event will answer for you. First, what does this cost in energy, and why can a company that meters

my tokens to the fraction of a cent not tell me the watt-hours—what is the silence protecting [20]? Second, where does my data go when the agent reads it, and does the documentation tell me, or does it change the subject to my productivity [5]? Third, what is the attack surface—can someone who can put text in front of my assistant issue it instructions, given that this is the single top-ranked threat in the field [27]? And fourth, who captures the value and who bears the cost—because the answer, on present evidence, is that the value is private and the costs are socialized.

None of these questions requires you to reject the tools. Broussard's technochauvinism has an equal and opposite error—reflexive refusal—and it is no more useful than reflexive faith [22]. The point is to see the artifact clearly enough to price it honestly: a genuinely capable next-token predictor, running on an industrial energy base the vendor won't disclose, ingesting your data through channels the guides won't foreground, exposed to a manipulation vector the security teams admit they can only mitigate, owned and gated by a platform whose business model is your dependence. That is a defensible thing to use with your eyes open. It is not an expert, and the moment you accept that it is, you have accepted the one claim engineered to stop you from asking anything else.

The enchantment, Crawford reminds us, works by making the method dazzling so that the purpose goes unexamined [22]. The purpose of the "PhD-level expert" line was never to describe GPT-5. It was to dress a power plant in a professor's gown and send it into your workflow before you thought to check what was running underneath. The literacy that matters this week—and every week the next model launches—is the habit of checking anyway.

[20] Tabla de tarifas de ChatGPT (precios Enterprise basados en tokens)

[5] Documentación de GitHub Copilot

[27] ¿Qué es un Ataque de Inyección de Prompt en IA? La Amenaza Oculta que ...

[22] Artificial Unintelligence - How Computers Misunderstand

[22] The Atlas of AI - Power, Politics, and the Planetary Costs

References

1. Amazon CodeWhisperer Documentation
2. Are the images generated by openai dalle commercially available and who ...
3. Assistance par intelligence artificielle pour les développeurs dans l ...
4. CodeWhisperer is becoming a part of Amazon Q Developer
5. Documentación de GitHub Copilot
6. Empoisonnement de données : la menace invisible qui cible les grands ...

7. Gemini Code Assist - Visual Studio Marketplace

8. Gemini Code Assist consumer accounts | Google for Developers

9. Gemini Code Assist overview | Google for Developers

10. Gérer et sécuriser les agents IA au sein de l'organisation - Cloud ...

11. Microsoft 365 Copilot adoption guide and overview for IT admins

12. Microsoft 365 Copilot adoption report | Microsoft Learn

13. Mises à niveau et limites des applications Gemini pour les abonnés ...

14. Modelos y precios para GitHub Copilot

15. No, ChatGPT Isn't Making You Dumber—What the Study Really Reveals

16. Overview of Power Platform and Copilot Studio reference ...

17. Risk and Safety Evaluators for Generative AI - Microsoft Foundry

18. Rollout Microsoft 365 Copilot to your organization

19. Safeguarding LLM security & safety evaluations | Microsoft Learn

20. Tabla de tarifas de ChatGPT (precios Enterprise basados en tokens)

21. Tarifario de ChatGPT (Business, Enterprise/Edu) - OpenAI Help Center

22. The Atlas of AI - Power, Politics, and the Planetary Costs

23. This AI notetaker won't sell surveillance to your boss

24. Unable to Deploy Anthropic Claude Models in Azure AI - Quota ...

25. Why voice actors are fighting Hollywood AI - Rest of World

26. ¿Qué es Microsoft 365 Copilot? | Microsoft Learn

27. ¿Qué es un Ataque de Inyección de Prompt en IA? La Amenaza Oculta que ...