

Phase Lock

A manifesto on brain computer interfaces and
preserving human agency in the age of AI.

SATURNIN PUGNET

October 2026

For the full interactive experience, read at phaselock.satpugnet.com

Phase Lock

A manifesto on brain-computer interfaces and preserving human agency in the age of AI.

SATURNIN PUGNET

phase lock /feɪz ˈlɒk/ *n.*

The exact threshold where two independent systems collapse into a single, unified frequency.

In **physics**, the moment an oscillator binds its rhythm to an external force to preserve structural stability. In **electronics**, a circuit anchoring itself to a master reference to eliminate drift. In **neuroscience**, the spontaneous alignment of firing neurons to an outside stimulus.

Humanity and AI are now about to enter *Phase Lock*.

Executive Summary

Reality is bending.

The future, which used to arrive in decades, now arrives by the month. Every day that passes feels faster and *stranger* than the previous one.

The intuitions our species built over millennia, at the slow pace of generations, are coming apart at a dizzying speed. Our economic utility, our role in society, and our very meaning in the universe are all being challenged; and as our philosophical anchors are *fracturing* one by one, we drift into the unknown.

Today, AIs write, code, reason, plan, and increasingly act on our behalf; and while humans are still in the loop, it feels like our presence is fading. As AI is gaining in agency by the day, one has to wonder: *can humans still control machines?*

At the center of this upheaval lies a growing mismatch between human cognition and machine intelligence, AI capability is racing ahead faster than human biology can keep up.

This is what I call **the bandwidth problem**.

As AI becomes more capable, our ability to direct, understand, and govern it does *not* scale with it. We have opened *Pandora's box* and we are attempting to steer what has emerged through keyboards,

screens, and clicks: interfaces designed for an earlier age and *unequal* to the task now before them.

Unless we develop scalable ways to align humans and machines, human agency will steadily recede from the systems shaping the world. *Either we align AI or we get left behind.*

There are two ways to do that: **alignment-by-control**, which constrains the machine from the outside, and **alignment-by-integration**, which couples human values and intents to the machine directly. Control can hold for the next few years, but ultimately, only integration will last.

The brief interval in which humanity can still determine the terms of its integration with AI before it grows too *powerful* to steer is the **Phase Lock Window**. Once it shuts, the choice is gone and humans are left permanently out of the loop, with no say in a future run by machine intelligence.

This manifesto explores a breakout technology that offers us a path to alignment-by-integration:

Brain-computer interface.

BCI is one of the primary levers we have for aligning AI with humans, serving three simultaneous functions: it provides an *architectural blueprint* for an aligned intelligence (the brain), it *supplies training signal* so that AI learns human intents and values directly rather than inferring them from our cognitive artifacts, and it *widens the bandwidth* between mind and machine, and ultimately of the mind itself, so that humans can keep pace with AI.

Even though it remains widely dismissed, BCI is approaching the same inflection point AI crossed roughly *a decade ago*. The *existence proof is already here*: researchers can decode speech, reconstruct mental imagery, and detect intent from raw brain activity. Before the decade is over, thoughts will turn into AI prompts, imagined scenes will materialize on screens, and robots will respond directly to mental commands; and a few years later, that connection will become bidirectional, making it possible to write thoughts or emotions directly to one's mind.

The question is no longer whether BCI can work, but how it will **scale**. The practical bet of this manifesto is that the *same scaling laws* that produced modern AI (better hardware, larger datasets and bigger models) apply just as well to brain decoding. By treating the brain as a vast, noisy, largely unlabeled dataset, AI can extract structure from signals that were previously too complex to decode; the **ChatGPT moment of the brain is now around the corner**.

The fastest path to execute this strategy will be through non-invasive BCI. Consumer devices can reach millions of people without surgery, generating the large and diverse neural datasets the field has been lacking. Consumer demand, paired with the growing value of neural data to frontier labs and robotics companies, will then finance further development, setting a *flywheel* in motion.

This manifesto maps that transition: how BCI will evolve over the next decade, the role it will play in the AI race, why neural data will become a critical asset for frontier labs, and how neurotechnology as a whole will help humans preserve *agency* and *oversight* over increasingly powerful machines.

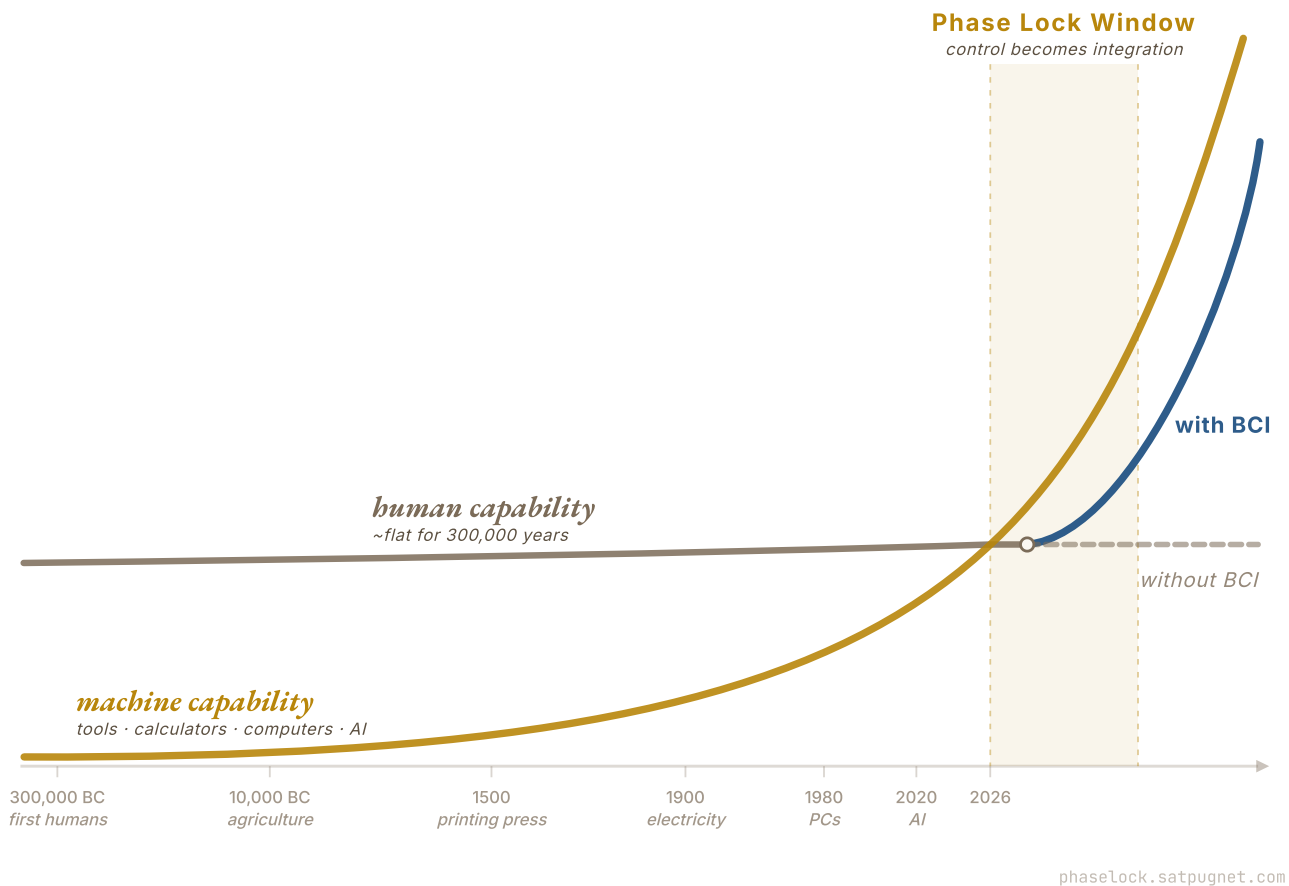
BCI and AI are now on a direct *collision course*. AI will give BCI its scaling playbook, its models, and its capital; and in turn, BCI will give AI a new training modality and a new interface.

From this mutually reinforcing dynamic will emerge a new lever for *controlling* and *aligning* AI with humanity.

AI will scale BCI, and BCI will scale AI.

How we shape their convergence over **the coming decade will define the centuries to come.**

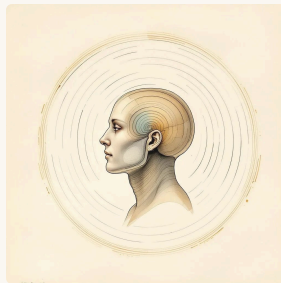
THE PHASE LOCK WINDOW



Human and machine capabilities over time. For 300,000 years, human capability set the pace; that era is ending. The Phase Lock Window is now open: the brief interval during which the relationship between human and machine intelligence can still be shaped before it moves beyond our control. BCI offers us a path forward through alignment-by-integration.

Outline

PART I	The Signal	6
	<i>Can the brain be read?</i>	
PART II	The Scaling	30
	<i>Does scaling work on the brain?</i>	
PART III	The Collision	68
	<i>Why does BCI matter for AI?</i>	
PART IV	The Aftermath	115
	<i>Implications for the future?</i>	



PART I.

The Signal

Can the brain be read?

01 What Is BCI?

“The brain is a world consisting of a number of unexplored continents and great stretches of unknown territory.” — Santiago Ramón y Cajal

To most, ChatGPT was a surprise. Three years before it landed, AGI¹ was a taboo topic, something people who read too much science fiction would bring up. Outside of a very niche subcommunity in SF and London, it was deemed absurd to think that within a few years a machine would surpass us at most tasks. The world expected decades if not centuries before it would be able to talk to an intelligent machine.

The world was wrong.

After a collective miss of that magnitude, intellectual humility is the least we can do. However, the right response is not just to ask why almost everyone failed to see LLMs coming, it is to ask ourselves: *what are we getting just as wrong today?* What transformational technology is on a clear exponential trajectory while being dismissed as science fiction by most?

Brain-computer interface is certainly a prime candidate. Too cryptic for the masses, seemingly impossible to build, and with an expected timeline of decades before anything meaningful comes out of it. All while the data is telling us a very different story.

Indeed, BCIs are about to surprise the world, and the ramifications of the technology will extend far beyond neuroscience, reaching all the way to frontier AI. An impact which will be vastly magnified by the collective ignorance of the impending arrival of the technology. Over the next five years, the field will move from an interesting research project to the center of the technological frontier.

BCI is now only a few years away from its ChatGPT moment.

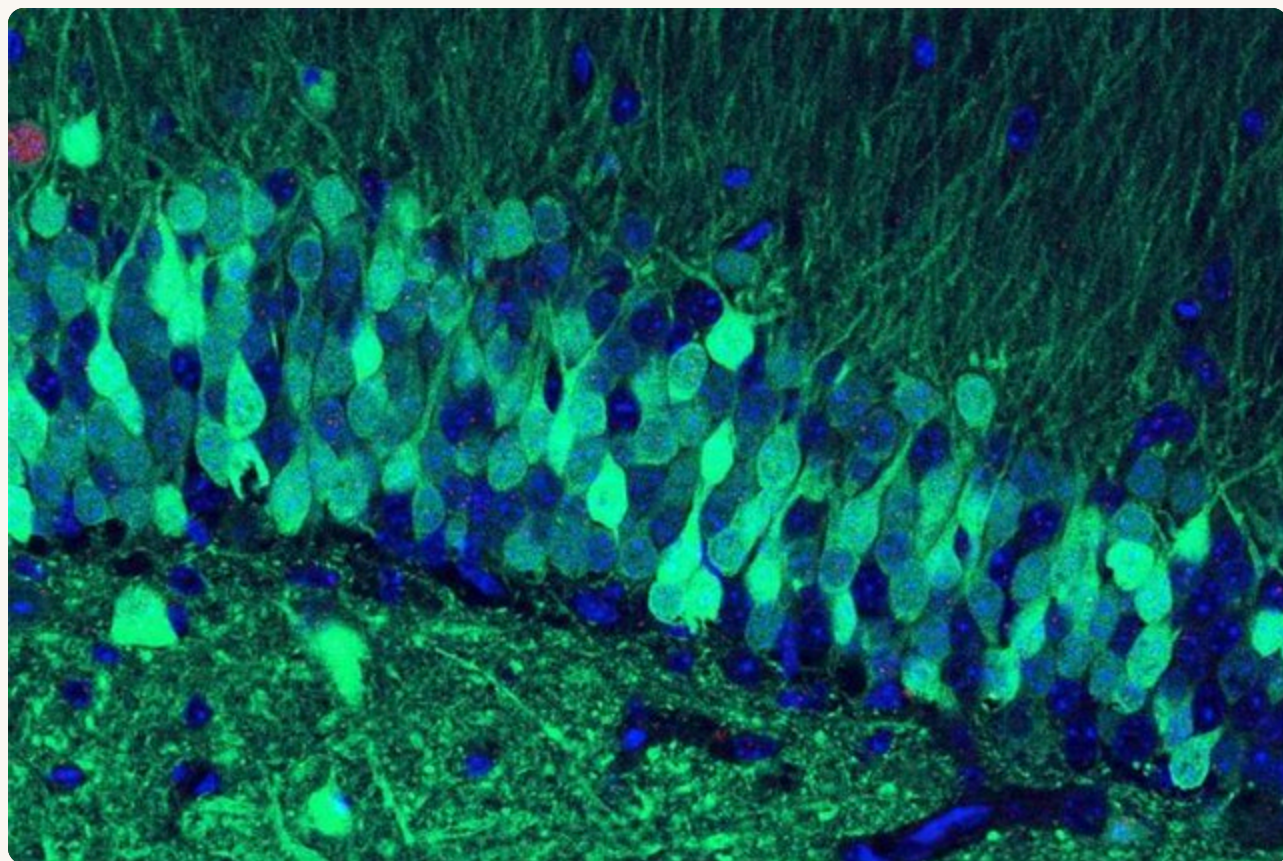
So what exactly is BCI? A brain-computer interface (or BCI) is exactly what the name suggests: a way for *the mind to communicate directly with a computer*. Instead of relying on hands to type or a voice to speak, BCI systems read the natural signals produced by the brain and translate them directly into digital commands. They remove the physical barriers between our thoughts and machines.

1. An Artificial General Intelligence defines an AI that is better than human across a broad range of cognitive tasks.

It may sound a lot like *magic* or *telepathy* (and in many regards it is), but remarkably, it is possible. We are already able to send words, images and emotions to machines via the power of thoughts only.

The deeper promise of this technology is a faster control of robots and digital systems and the removal of layers of *compression* between thought and expression; unlocking richer, **higher-fidelity communication between minds**, biological or artificial.

As simple as it is to describe, building BCI is in fact a *herculean* task, under this simple premise lives a wealth of complexity. Humanity has been trying to crack this problem for more than 50 years and just like AI, it has been making slow, but compounding progress over the decades.



Memory cells in a mouse brain. Somewhere in your head, a clump of cells like this is storing this exact sentence; BCI is the attempt to decode it. Source: [NIH](#).

Three Paths

When people say “BCI” they tend to mean three quite different, yet adjacent, technologies. The simplest way I have found to explain the difference is the stadium analogy.

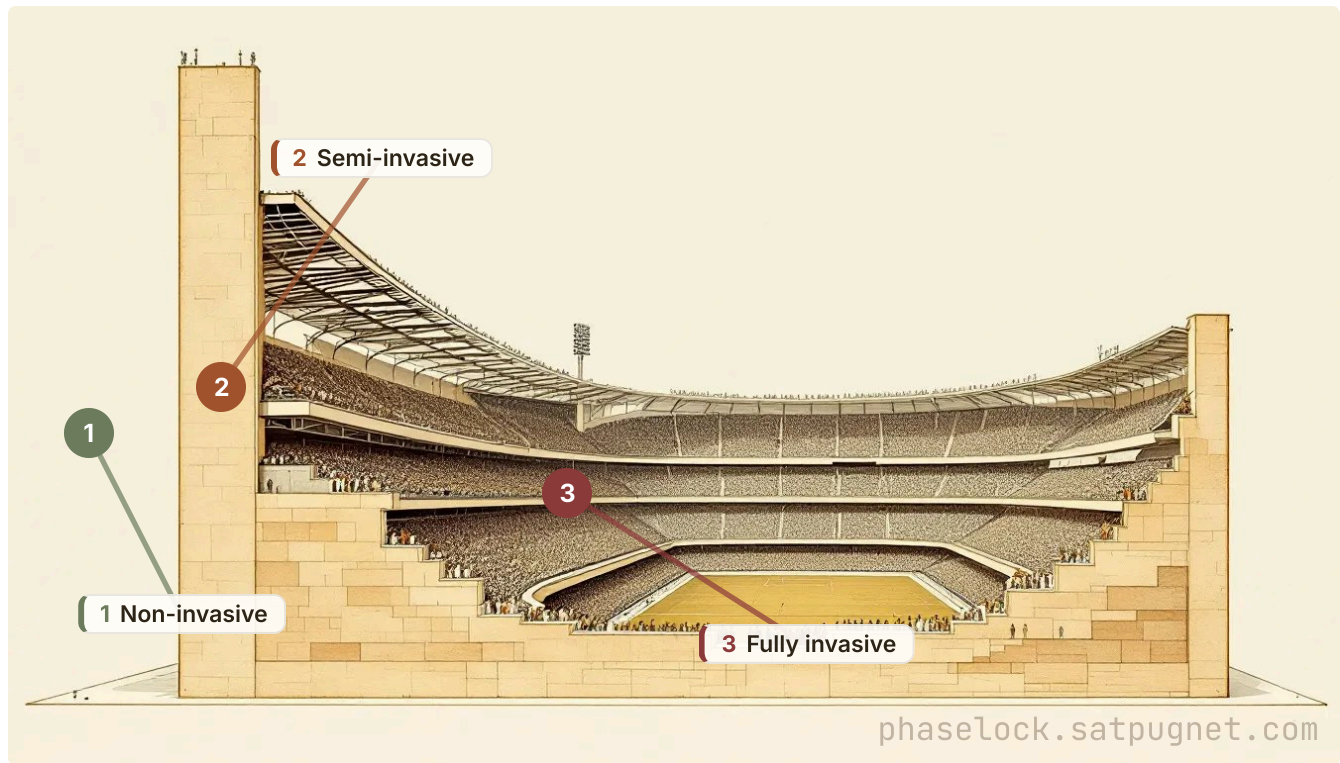
Imagine you are standing outside a packed stadium during a football match. The crowd is screaming. You want to know what they are saying.

You can stand outside the stadium and hold a microphone up to the wall. You will hear something; it will be muffled and washed out and dominated by the loudest, simplest patterns. That is **non-invasive** BCI. The brain equivalent is sensors sitting on the scalp and picking up electrical, magnetic, or hemodynamic activity through the skull. No surgery, and no FDA approval is required for consumer applications, it has the broadest theoretical user base, but also the noisiest data.

Alternatively, you can drill a hole in the wall and put a microphone right against the inside surface. The signal gets dramatically cleaner, although you still cannot quite pick out individual voices. That is **semi-invasive**; devices placed under the skull but not through the dura mater. Comparable to outpatient procedures we already do at scale (LASIK, cochlear implants, pacemakers). This path takes regulatory work, and recovery rooms.

Or you can walk into the stadium and put a microphone next to a specific person's mouth. That is **fully invasive**, electrodes directly in brain tissue (what [Neuralink](#) is doing). The cleanest signals possible but also the longest road; involving surgery, regulations, and generational timescales for adoption.

STADIUM ANALOGY



WHAT EACH MIC HEARS



1 Non-invasive
muffled, washed out



2 Semi-invasive
clearer, still blurred



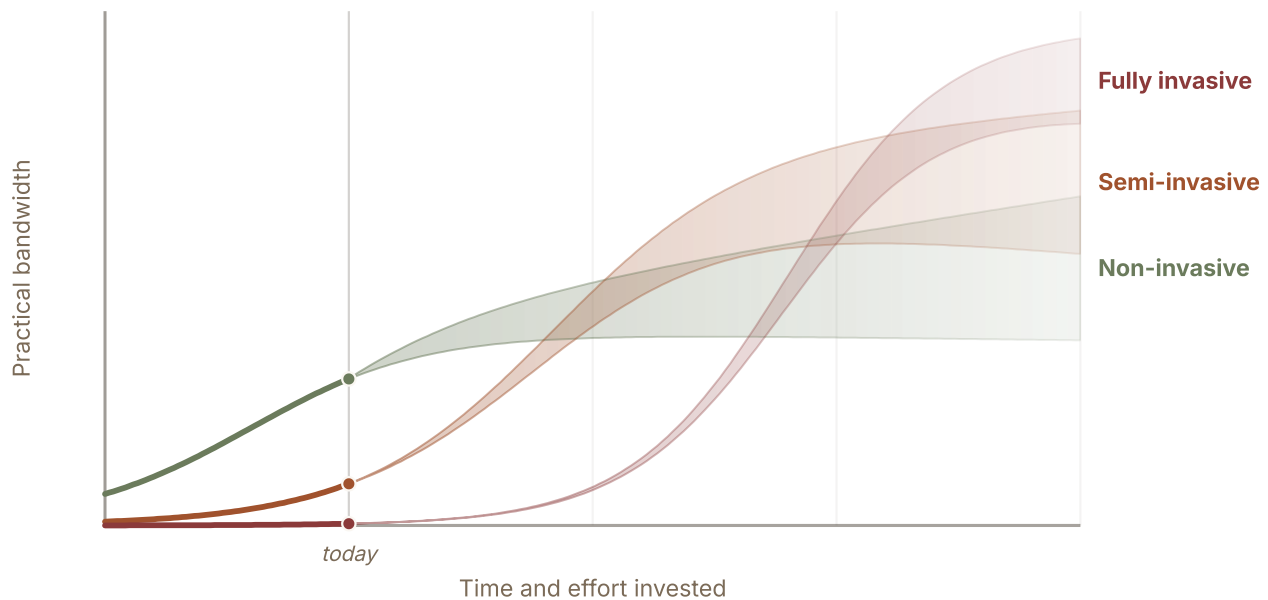
3 Fully invasive
clean, individual voices

Three places to put a microphone for a screaming stadium, each illustrates a different BCI strategy.

If you track how these three paths improve over time and effort, they each form an S-curve. Progress starts slow, accelerates rapidly as the technology matures, and eventually flattens out when it hits the physical ceiling. The metric that matters most is practical bandwidth: how much useful information we can actually pull out of the head and put to use at scale.

Non-invasive has the capacity to scale faster over the next few years and reach a viable threshold of market value because it does not require a surgeon, but its capability ceiling might arrive early because the skull blocks signals. Fully invasive has a [theoretically much higher ceiling](#) ▶, but it must first crawl under the heavy floor of clinical regulation and surgical throughput before it can even start to scale.

THREE S-CURVES OF CAPABILITY



Non-invasive advances first and scales fastest, though signal attenuation through the scalp and skull constrains its maximum resolution. Fully invasive crawls under the regulatory floor for years but eventually reaches the highest ceiling.

The important point is that nobody actually knows where the ceiling for any of these curves sits today. We have not climbed far enough to hit the head-room of the possible, until we do, the different modalities are just different ways of asking the same question: What is the fastest way to decode the mind at scale?

SIDENOTE

There is one more distinction worth drawing. **Reading vs. writing.** Reading the brain (pulling structured information out) is a fundamentally easier problem than writing to the brain (putting an intention or experience in). Sensing the magnetic, electrical, or hemodynamic shadow of activity through the skull is hard but tractable. Stimulating a specific neuron or population from outside the head, with the precision to *cause* a thought rather than disrupt one, is a much harder problem. It may simply not be possible non-invasively at high fidelity.



Wilder Penfield, 1934. One of the fathers of modern neuroscience.

But philosophy does not build hardware, what matters is the data. So let's look at what non-invasive BCI can actually do today, and establish the physical baseline of what is possible.

02 The Art of the Possible

“When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.”

— Arthur C. Clarke, First Law, 1962

BCI capabilities are already beyond what most people expect. Here are four things that happened in the last three years, which you probably did not hear about.

[A team at UT Austin](#) ▶ placed a person in an fMRI scanner, played them a podcast, and **reconstructed the sentences out of brain activity**, sentence by sentence; not the audio but the *meaning* itself. They can do this for silent inner narration too: you think the sentence, and the model writes a paraphrase of it.²

TANG: HEARD VS. DECODED

Actual stimulus

i got up from the air mattress and pressed my face against the glass of the bedroom window expecting to see eyes staring back at me but instead finding only darkness

i didn't know whether to scream cry or run away instead i said leave me alone i don't need your help adam disappeared and i cleaned up alone crying

that night i went upstairs to what had been our bedroom and not knowing what else to do i turned out the lights and lay down on the floor

i don't have my driver's license yet and i just jumped out right when i needed to and she says well why don't you come back to my house and i'll give you a ride i say ok

Decoded stimulus

i just continued to walk up to the window and open the glass i stood on my toes and peered out i didn't see anything and looked up again i saw nothing

started to scream and cry and then she just said i told you to leave me alone you can't hurt me anymore i'm sorry and then he stormed off i thought he had left i started to cry

we got back to my dorm room i had no idea where my bed was i just assumed i would sleep on it but instead i lay down on the floor

she is not ready she has not even started to learn to drive yet i had to push her out of the car i said we will take her home now and she agreed

Exact

Gist

Error

Left: what the subject heard. Right: what the model decoded from their fMRI activity, word by word. Color-coded for exact / gist / error.

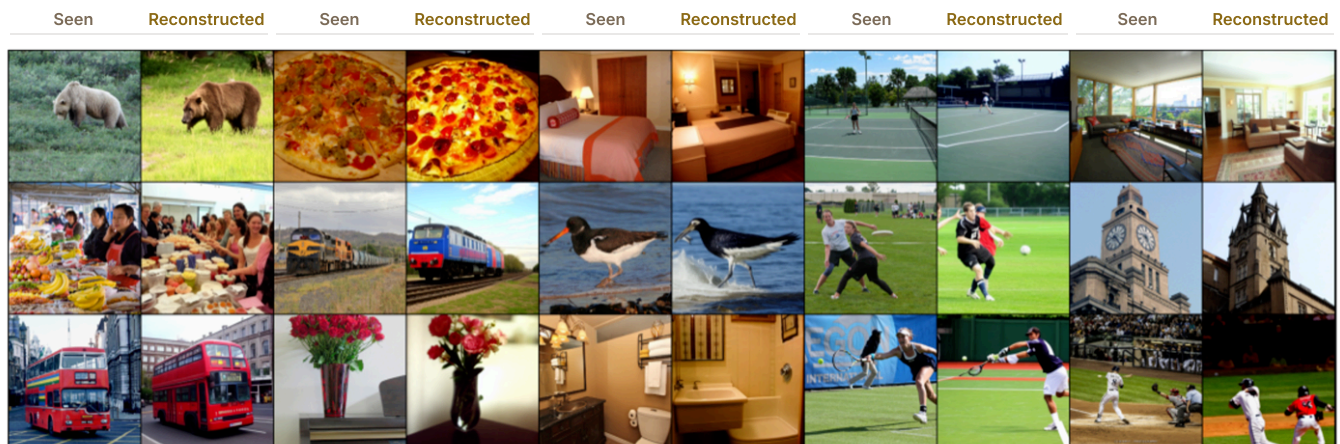
Source: [Tang et al., Nature Neuroscience 2023](#) ▶.

[A team at Stanford](#) ▶ implanted microelectrode arrays in a paralyzed person's motor cortex, asked them to imagine handwriting letters with their hand, and **decoded typed text at 90 cpm**. That is faster than most of us type on a phone. ([A team at UCSF](#) ▶ reached comparable rates a couple of years later, decoding intended speech directly from the cortex.)

[A team at Princeton](#) ▶ captured fMRI activity from someone looking at a photo and **reconstructed that photo from the brain signal alone**, using diffusion models trained on neural-image pairs.

² Tang's work stands on the shoulders of the Gallant Lab at Berkeley, who built the first cortical semantic maps a decade earlier; [Huth et al. 2016](#) ▶, [Nishimoto et al. 2011](#) ▶, and others laid the foundations that made it possible.

MIND EYE: IMAGES SEEN VS. RECONSTRUCTED



In each pair, the left image is the photograph the subject saw, while the right was reconstructed solely from brain activity, without help from the photo. Source: [Scotti et al., MindEye 2023](#) ▶.

[A team at Tongji](#) ▶ extended the same idea from stills to video. They played people short clips, and then **decoded the videos from their minds**. Pull that thread forward a decade, and you can imagine your phone displaying live visuals of what you just imagined.

NEUROCLIPS: WATCHED VS. RECONSTRUCTED

AIRPLANE

WATCHED



RECONSTRUCTED



STREET

WATCHED



RECONSTRUCTED



What the subject watched, and what the model reconstructed from their brain activity alone (using fMRI). Source: [Gong et al., NeurIPS 2024](#) 🐦.

We extracted a video out of someone's head from only a few tens of hours of scan data, *how is that even possible?* As shocking as it may sound, this is the current state of the field.

And while none of these demonstrations, by itself, proves that scaling the technology will be easy (fMRI does not scale to the living room and invasive microelectrode arrays do not scale to healthy consumers), they show something much more fundamental: that the signal is there, and that modern models can read it when the data is clean enough. And this is not just wishful thinking anymore, those are published, peer-reviewed demonstrations. Together, they all point to the same structural truth: **if a thought leaves a trace in the signal, a large enough model can find it.**

The remaining question is how much of that signal can be recovered non-invasively, at scale, in real life.

As can be seen from those examples, the current frontier work in BCI is doing impressive things in narrow domains with tiny datasets. **One of the core claims of this manifesto is that the BCI field is about to undergo the exact same phase transition AI went through in 2017: moving from clever human engineering to brute-force scaling.** The same structural shift which produced LLMs.

But before we look into the most promising way to scale BCI, there is one important piece of human physiology worth discussing in the context of human-machine communication. It sets both the *floor* a BCI must surpass to become useful and the natural *ceiling* on our ability to communicate with AI; I am talking about **our peripheral nervous system.**

03 The Peripheral Nervous System Bar

“The hand is the cutting edge of the mind.” — Jacob Bronowski, The Ascent of Man, 1973

The **peripheral nervous system** is key to understanding the role and limitations of BCI in interfacing with machines.

The fancy wording “peripheral nervous system” simply refers to the biological hardware that connects your brain to the outside world, otherwise known as eyes, ears, hands, vocal cords etc. The PNS is the way your brain interacts with reality; any effect your mind has ever had on the world has first had to pass through it.

It occupies a strange position between humans and machines. For AI, it is a *painfully narrow bottleneck*; while for BCI, it is an *extraordinarily capable incumbent*.

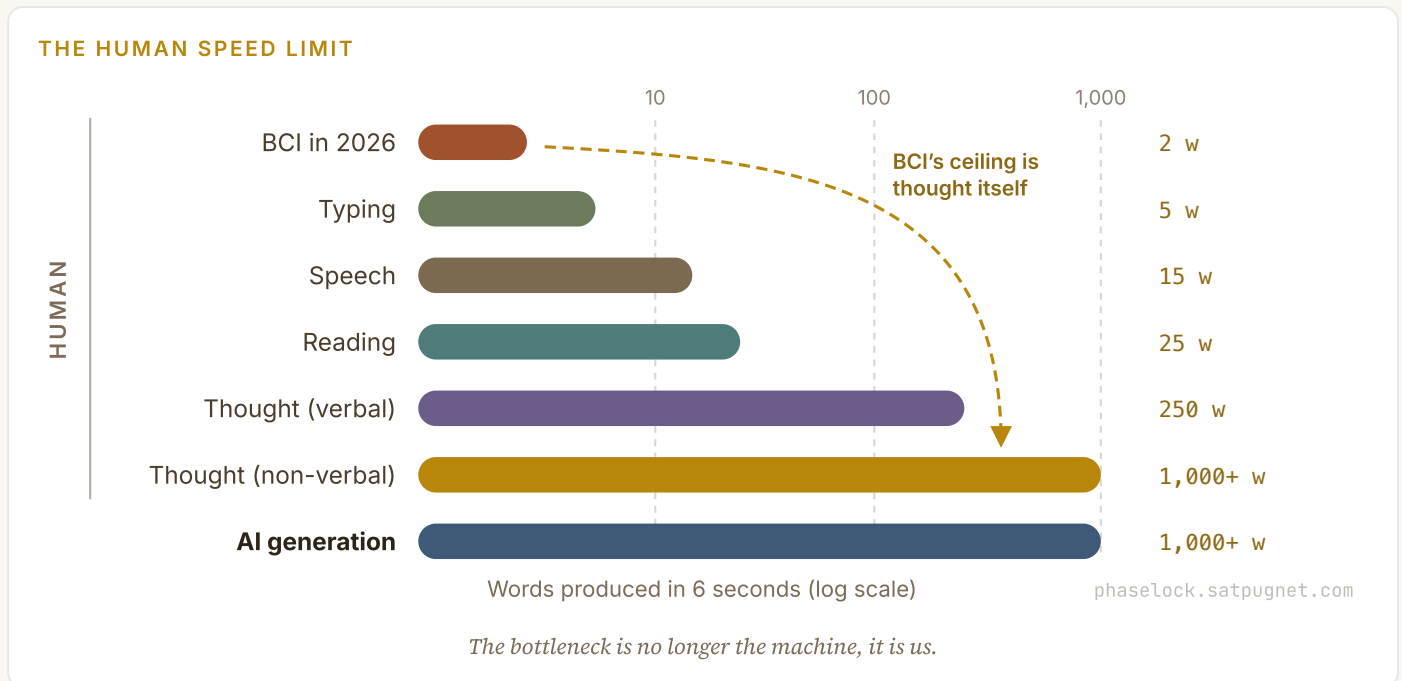
Too slow for AI

From the perspective of AI, your peripheral nervous system is slow, too slow, and is the first half of the bandwidth problem³ (the other half being the speed of thinking itself).

Indeed, every computer interface humans have ever built sits between two clocks (a clock in information theory defines how quickly a system processes information): the biological one and the machine one; and humans now have a slower clock speed than AI. At this point, the human clock is inconveniently slow: speech runs at about 150 words per minute, typing at 50, reading at maybe 250, inner thought is harder to measure but probably runs at 1,000 to 4,000 words per minute equivalent,

3. The problem arising from the fact that AI can now take in and produce information thousands of times faster than a human can ingest and process it.

and a great deal more if you count the non-verbal parts of cognition. All while AI cognition is orders of magnitude faster.



For most of computing's history, the gap between the two clocks did not matter much. Machines were slow enough that a keyboard could keep up, and humans were comfortably able to control the machines; but AI changed that. **Machine capability has scaled by orders of magnitude in the last five years; yet the interface still routes through a keyboard.** Models can now read a book in seconds and write one in minutes, and they are only getting faster. Increasingly, the slowest component in the loop has become the human. Both our cognition speed and our interface bandwidth are becoming the *bottleneck*.

To provide an analogy, imagine trying to direct a civilization that experiences ten years between each word you speak. A ten-word instruction would take it a century to receive. That is the emerging human-AI interface problem: the machine can race through vast amounts of reasoning while human reasoning is left in the dust and intent still trickles in through speech, typing, and clicks.

That is where neurotechnology comes into play, it can help close that gap. As AI gets better at thinking and worse at waiting, the value of letting humans co-think with it grows fast; a direct neural channel removes the need to compress multi-dimensional thoughts and emotions into lossy, sequential streams of typed or spoken words. It lets you communicate your intent unfiltered to the AI, and, at the limit, it lets you upgrade your cognitive hardware to match the machines.

SIDENOTE

Zooming out, this mismatch stems from the fact that **two forms of intelligence are now co-evolving**; machine intelligence, scaling at the rate of compute, and human intelligence, scaling at the rate of biology. Learning to co-exist with a new form of intelligence which surpasses us on so many levels is an unprecedented challenge. Each year AI beats most of the new benchmarks we create for it, it unlocks new skills, and it improves itself at an accelerating rate. Figuring out ways to successfully navigate this transition while keeping humans in the loop is a fundamentally unresolved problem, and a hard one too.

BCI is promising to be a critical tool to address this growing AI-human mismatch. It is arguably **the only tool we have that focuses on raising human capability rather than containing AI capability**. Currently, most of the levers the alignment⁴ community is reaching for (interpretability, evals, governance, scalable oversight) are defensive; they try to either slow AI capability, constrain it, control it, or interpret it after the fact⁵. And in the long run, only a lever that raises the human capability can keep the human side of the loop meaningful as the machine side continues to improve. **That is what makes BCI, and neurotechnology more broadly, a fundamental pillar of any long-term arrangement in which humans stay in the loop.**⁶

We will come back to the topic of AI and BCI convergence in the second half of this manifesto ([Part III](#) ► and [Part IV](#) ►). But before doing so, let us continue our exploration of BCI.

Too fast for BCI

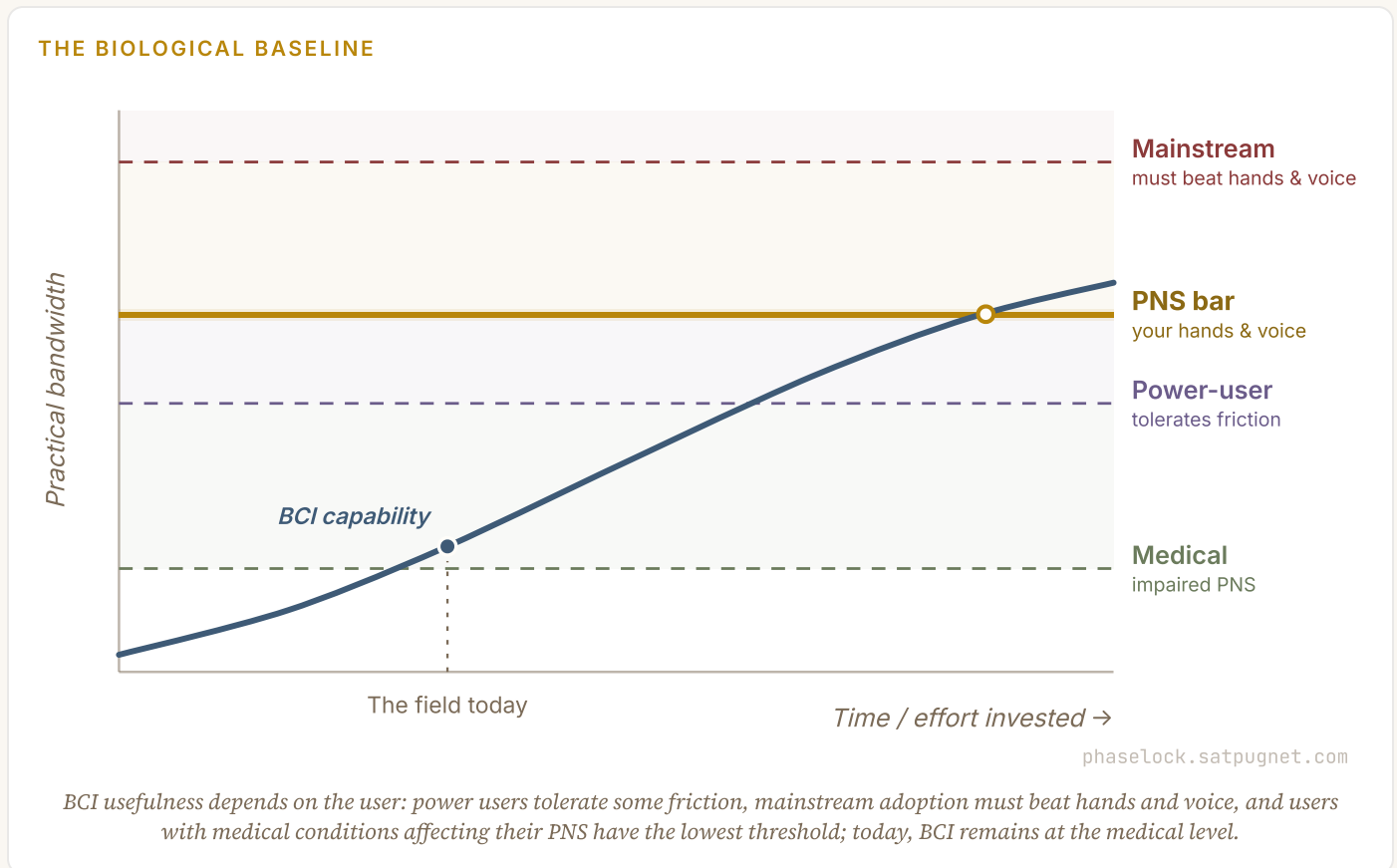
From the perspective of BCI, your fingers can hit a key in a few hundred milliseconds with high accuracy, and your tongue and jaw produce roughly 150 wpm of speech. The bar a BCI must clear to be useful to a person whose hands and mouth still work is therefore extreme. It is not enough for a BCI to read your intent, it has to read your intent with an inference latency low enough, and an

4. Alignment, in AI safety, is the problem of ensuring an AI system's behavior matches what humans actually want, rather than a misleading proxy of it. The harder version, often called *scalable alignment*, asks how we ensure this alignment continues as AI becomes smarter than us. This is an unsolved problem.

5. I do not discount the current work in interpretability, evaluations, governance, scalable oversight etc... Those efforts remain absolutely critical. The point I am making is rather that neurotechnology has an equally important role to play in alignment, complementary to these approaches.

6. To be honest, if AI were not advancing so quickly, I would advocate slowing down the development of neurotechnology as this is a very risky technology. But with AGI now so close, it would be a mistake to do so.

accuracy high enough, to *beat* what your hands and mouth already do for free. If sub-vocal text takes 800 ms to decode at a rate of 10 wpm, the user will simply use their thumbs.



This is why every reasonable timeline for BCI splits into two milestones. There is “does decoding work at all”, which we are largely past for many tasks; and there is “does decoding work at a level a person without an existing peripheral-nervous-system disability would pay for”. Here, progress is non-linear; small fidelity gains near the threshold drive disproportionate usefulness. Voice recognition is the cleanest analogy, a 95% transcription system is unusable, whereas a 99.9% system can be used everywhere.

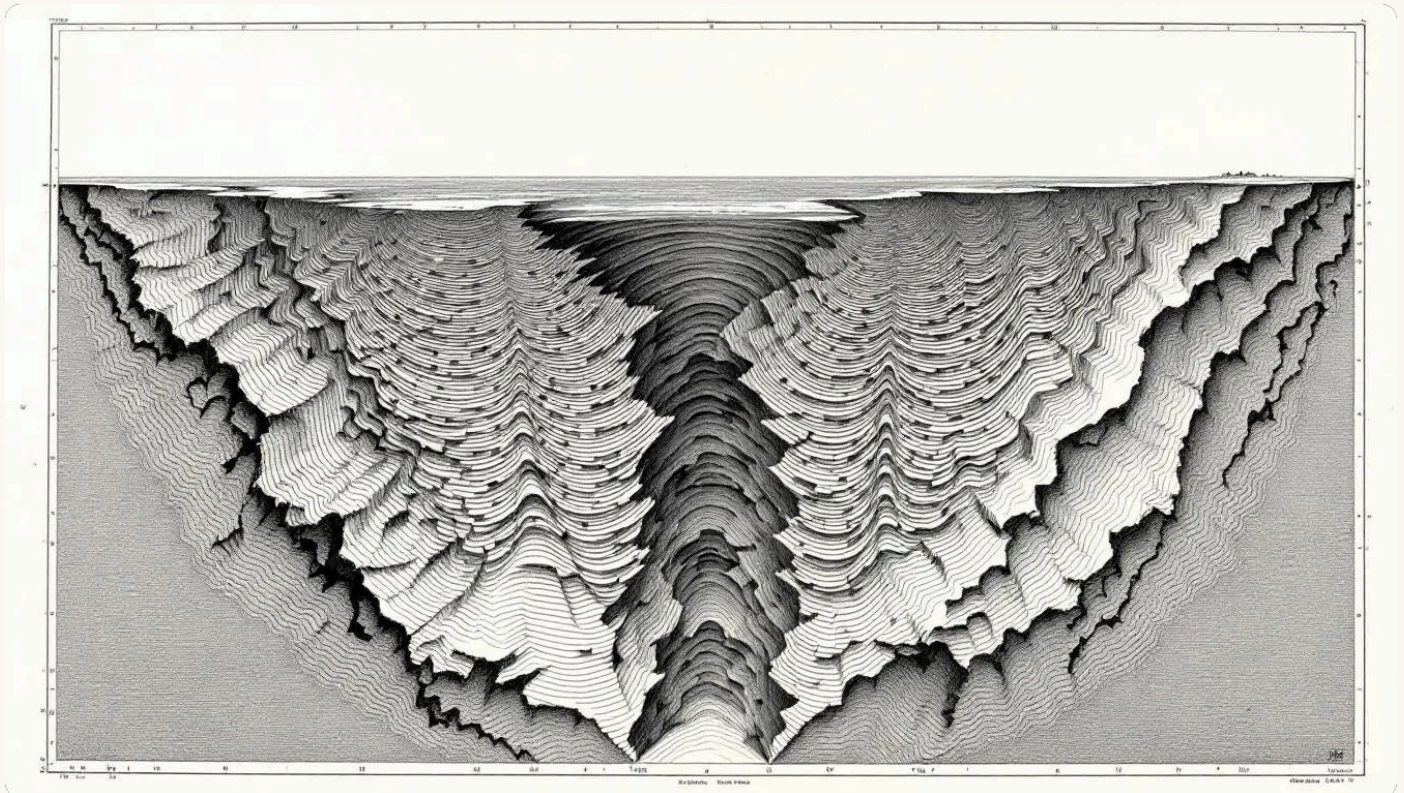
This is also why **medical applications have the lowest bar to clear, and are the first to fall**. Patients with severe neuromuscular conditions (ALS, locked-in syndrome, paralysis) do not have an intact peripheral nervous system to compete against. The bar to do something useful for them sits well below the consumer bar. Tolerance for physical risk is higher, and the urgency is absolute. Almost every working BCI demonstration to date has been in this regime.

But medical only reaches a few thousand patients, the only way to reach hundreds of millions, at least in the short term, is through a consumer non-invasive product. Non-invasive is the fastest path to scale.

04 The Fastest Path to Scale

So far, I have had a stronger focus on non-invasive BCI, as opposed to invasive or semi-invasive. There is a reason for that. Non-invasive BCI has a couple of advantages over alternatives which make it a much more suitable candidate for applying the AI scale paradigm. And while **the cleanest signal we can get comes from inside the brain, the fastest path to scale BCI will be through non-invasive.**

Twenty years ago, the consensus held that meaningful brain reading required going *inside the skull*. The signal-to-noise through bone seemed too poor for anything else. That belief is why the surviving old-guard BCI companies ([Neuralink](#), [Blackrock](#), [Synchron](#), [Paradromics](#), [Precision Neuroscience](#)) all chose surgery from the start; the bet was reasonable at the time. **For writing back into the brain, it still is**, putting precise patterns of activity into specific neural circuits needs resolution only invasive or semi-invasive can deliver, and that still holds true this decade. **For reading, however, large language models have rewritten the rules of the game.** Models trained on enough biological noise can pull structure the unaided eye could not see; the cost of extracting cognitive signal from messy non-invasive signal has been dropping by orders of magnitude. The skull still attenuates the signal, but large models just get better at reading through it over time.

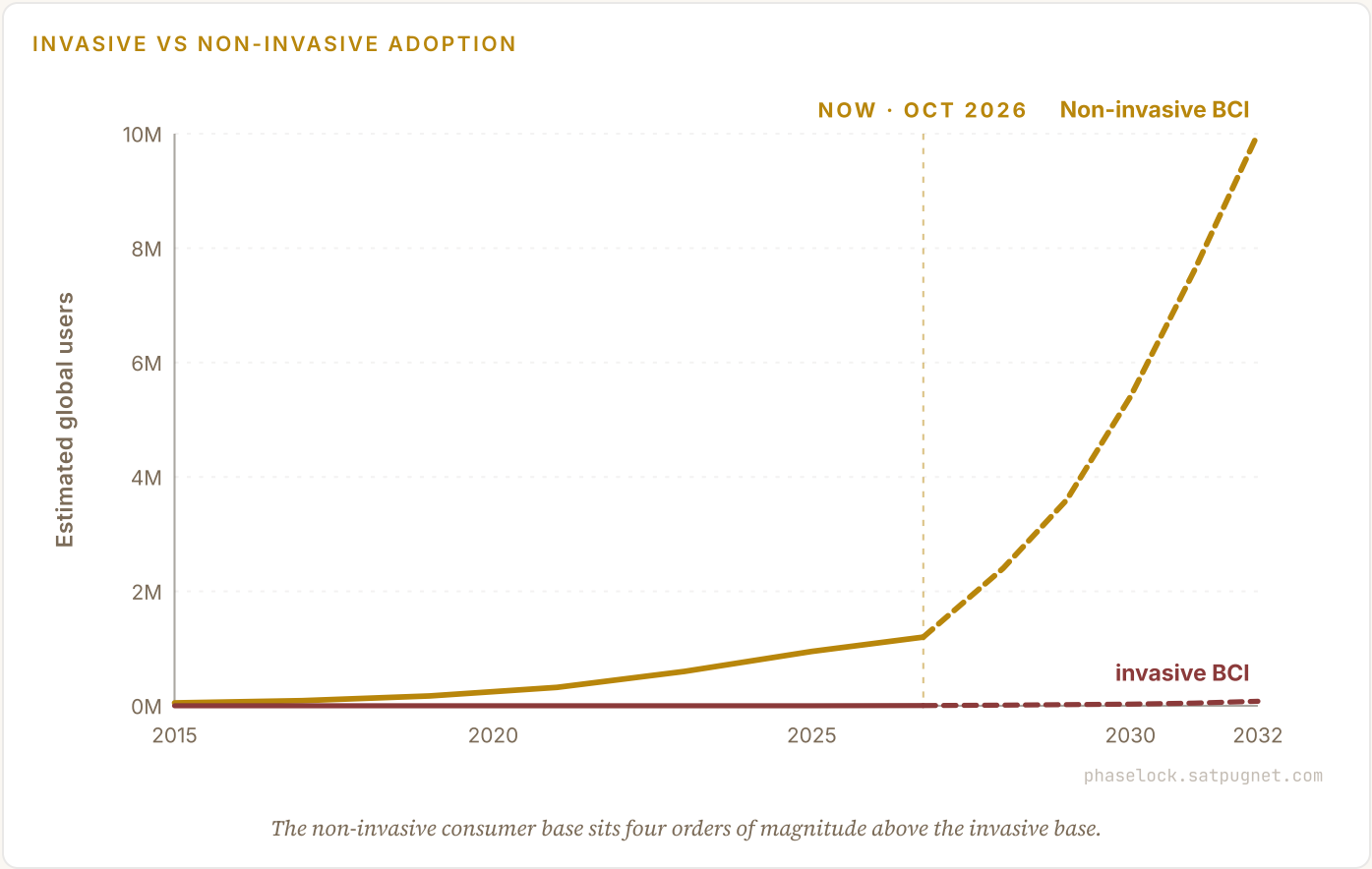


Sonar technology uses the same principle as some BCI devices, sending a pulse and then capturing its interaction with the environment allowing it to see otherwise invisible structures.

But if invasive signals are obviously the cleanest, why bother with non-invasive ones at all? *Why not just, as a field, commit to the invasive path?*

There are a few reasons.

First, surgery is the hardest constraint to overcome. On top of the fact that most healthy people are not exactly lining up for elective brain surgery. Brain surgery means *FDA pre-market approval* (a regulatory pathway that takes 7 to 10 years and tens of millions per device). It means *hospital procurement* cycles (18 to 36 months from outreach to first scheduled patient). It means *risk-averse insurance underwriting*. It means an *operating-room throughput limit*. Even at maximum velocity, a single neurosurgical center can only do dozens to low hundreds of implant procedures per year⁷. And even in a best-case scenario, a brilliant invasive technology working perfectly will only be able to reach a few thousand patients in the time a successful non-invasive product could reach tens of millions.



As of May 2026, fewer than 100 people worldwide had received chronic invasive BCIs, mostly through Neuralink and Blackrock/BrainGate, with annual growth still measured in dozens. Non-invasive

7. The entire US neurosurgeon capacity (roughly 4,000–5,000 practicing surgeons) caps invasive throughput at roughly 25,000 to 100,000 implants per year, and that is *even* if we assume partial robotic-surgery autonomy. Reaching one million healthy adults requires a fundamentally different surgical pipeline that simply will not happen this decade. The earliest plausible mass-scale invasive timeline lands in the late 2030s at best, with the broader build-out running deep into the 2040s.

consumer BCIs, by contrast, already exceed one million users. Muse alone reports over 500,000, Emotiv over 70,000, and BrainCo tens of thousands, with Neurable, NextSense, and Meta expanding the category.

Fully invasive systems will keep advancing, and they will eventually produce the highest-bandwidth channel into the brain. There is no doubt that they are the *final destination* of the BCI field, but physics, biology, and the medical establishment dictate their timeline.

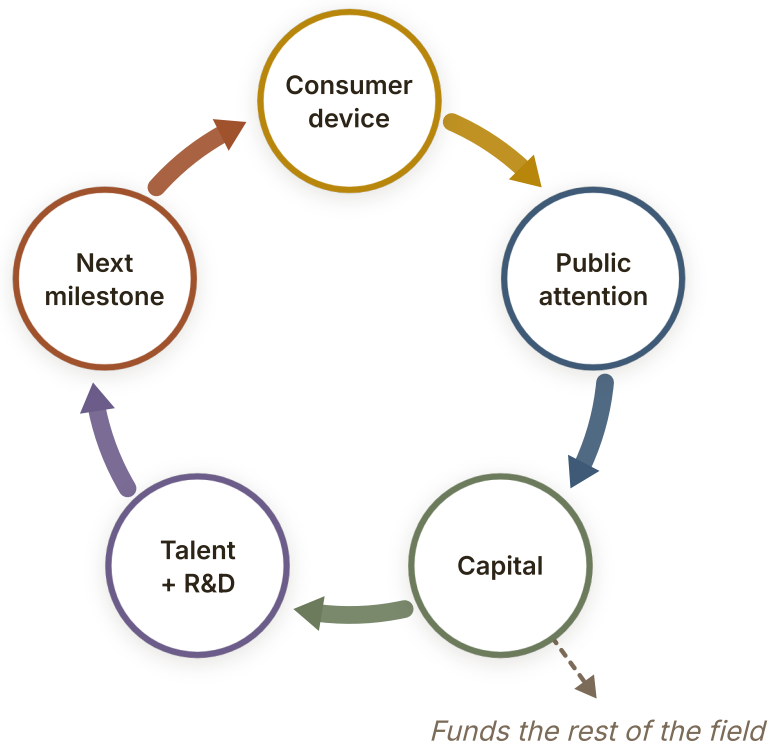
Second, we do not know the limit of non-invasive. While we know that invasive BCI is ultimately better in the long run, we have no idea what the performance ceiling is for non-invasive. The honest scientific position on non-invasive BCI is that we do not know how good it can get, because nobody has stress-tested it at scale. Every reasonable counterargument to non-invasive points to signal quality through the skull, but every reasonable counterargument has the same hole: it argues from a theoretical physics standpoint about a problem that is inherently empirical. No data-backed argument has shown that building a consumer device with non-invasive BCI is not achievable. The few experiments that have looked harder *show that scaling laws are still holding*.

Third, a working non-invasive BCI would cause a flywheel effect. The fastest way to accelerate any frontier technology is to ship a consumer product at scale. In the AI world, that was ChatGPT. A first working product of a groundbreaking technology ignites a flywheel, it generates attention, attention pulls capital, capital pulls talent, and talent builds the next iteration. But flywheels need a first push to start turning, **right now, the BCI flywheel is almost static.** There are some medical demonstrations and some lab papers, but no real consumer device. And the only chance to get there that we have within the next five years is *through non-invasive*.

Take Tesla. Before the Roadster⁸, electric cars were a fringe idea; a sliver of attention with almost no capital. Then Tesla shipped a flawed but working sports car, and the world saw, for the first time, that an electric car was a thing you could actually own and drive. Public attention shifted and capital followed. The trillion-dollar pipeline that built the Model S and the decade of EV companies behind it was only possible once that first basic car proved the category existed. **Non-invasive BCI is where Tesla's first car for BCI is made.** What will fund the rest of the field is then the public-belief shift that arrives when a working consumer device exists.

8. Tesla's first car.

THE CAPITAL FLYWHEEL



phaseLock.satpugnet.com

Once the flywheel starts spinning, BCI will accelerate dramatically and fund the rest of the neurotechnology field, the hardest part is the cold start.

SIDENOTE

But what if the skull is simply too thick? If you remain skeptical that non-invasive hardware can clear the consumer bar, and think it is a waste of effort, you can evaluate the worst-case scenario, the strategy is hedged.

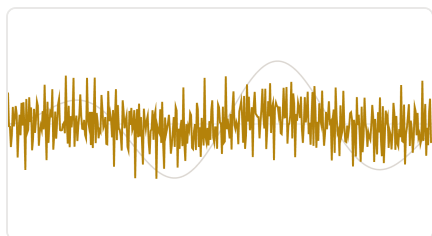
First, the engineering transfers. The decoder architectures, data conventions, and multi-modal sensor stacks being built right now are largely hardware-agnostic; they port cleanly to semi-invasive and invasive systems. The talent and capital that enter the field stay in the field.

Second, the data is an asset. Even a stalled consumer push can generate a large data corpus of multi-modal neural data; this would still be an extraordinarily valuable resource for the rest of the field. Even in the worst-case scenario, pushing the non-invasive limit would still accelerate the rest of the field.

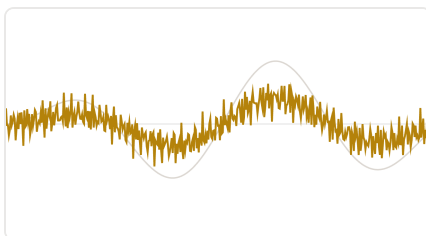
The BCI flywheel assumes it is possible in the first place to build a working product from non-invasive technology in the short term. **The bet is that the brain emits enough signal through the skull to decode useful information non-invasively.** The question is whether the signal that survives (through the skull, through movement, through a Monday afternoon, through a different user) carries *enough information* for better sensors and larger models to make sense of it at scale.

READING SIGNAL OUT OF NOISE

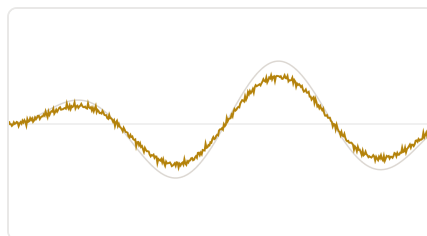
BURIED · 15% SIGNAL-TO-NOISE



BORDERLINE · 40% SIGNAL-TO-NOISE

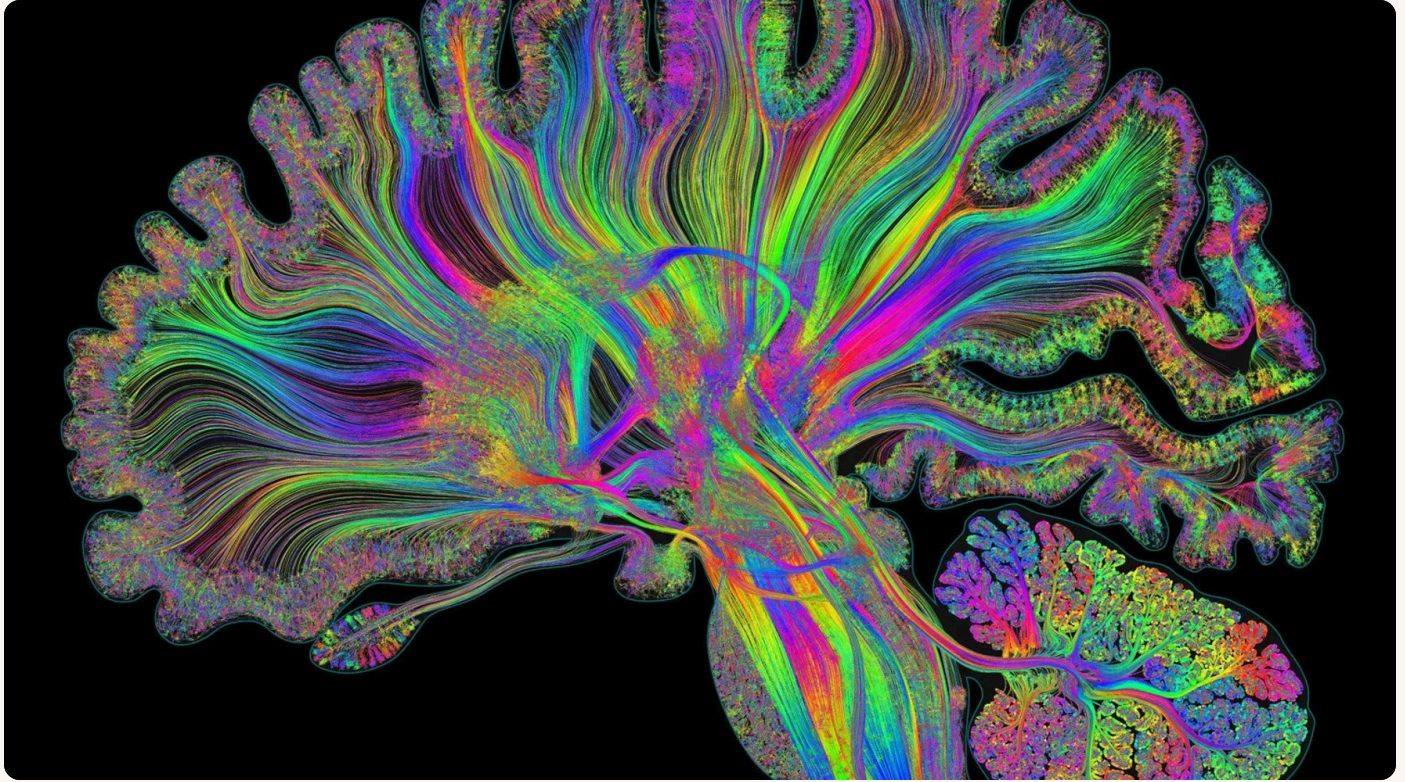


DECODABLE · 75% SIGNAL-TO-NOISE



Below a signal-to-noise floor the brain would be unreadable; the hypothesis is that enough signal survives to reconstruct the thought. This visual illustrates how increasing the signal-to-noise ratio makes the signal increasingly more decodable.

The fundamental thesis is not just that signal exists, it is that enough of that signal is stable across people, sessions, and contexts that pretraining can amortize the cost of learning one new brain. That is more of an *empirical* question than a *theoretical* one. We do not know the limit of non-invasive brain decoding yet because nobody has stress-tested it at scale. We will only know what the physics actually allows once we build the infrastructure to test it. **As of 2026, the true ceiling of the non-invasive hardware remains completely uncharted.**



86 billion biological wires, the tangled mess we are trying to decode.

This sets the threshold of capabilities that a consumer BCI must meet to be useful and scale. For the first time, the conditions to clear this threshold are coming together; and they draw a pattern that *I have seen before*.

05 Why Now?

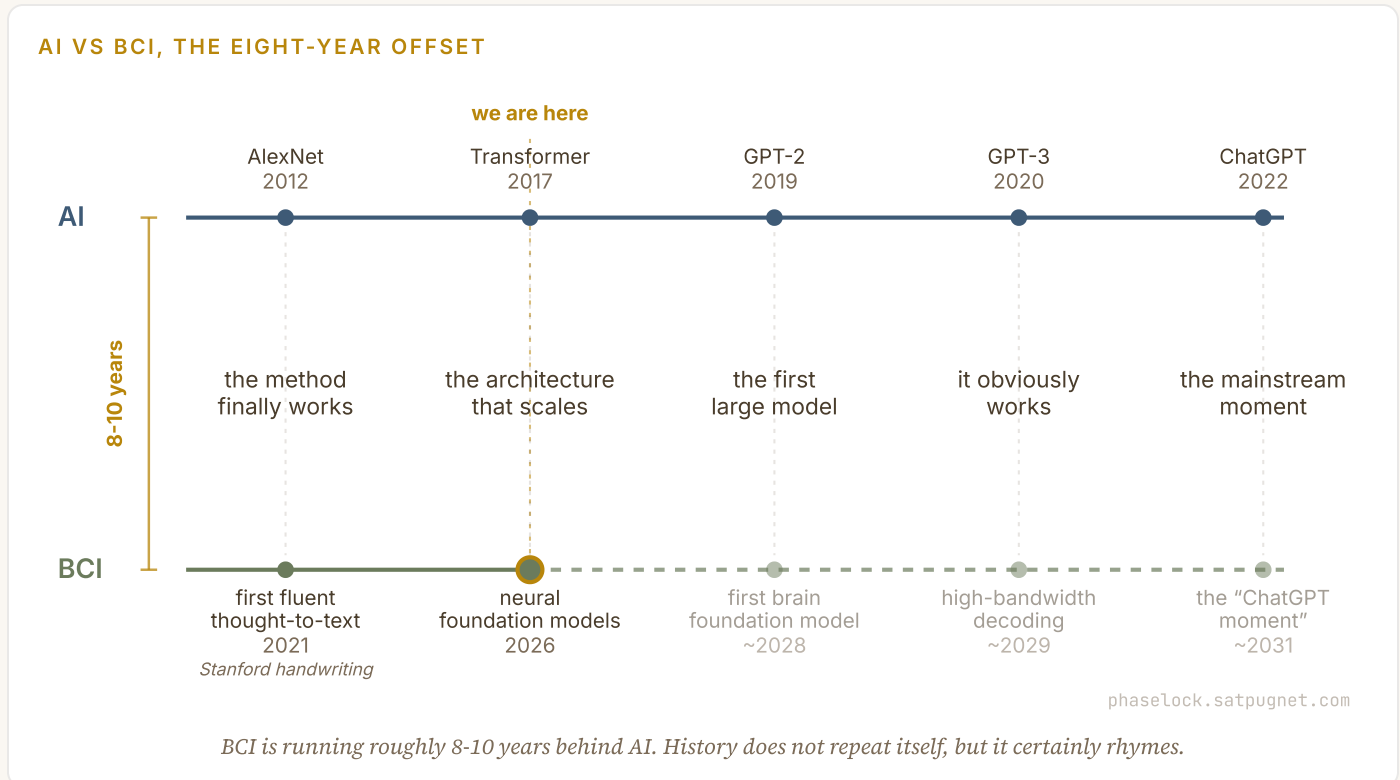
Back in the rainy London winter of 2018, I read a book which *changed my life*: [Life 3.0](#)⁹. I picked it up out of curiosity, but by the time I turned the final page, I was convinced that the transition towards AGI was already underway. From there onwards, I decided to *pivot* my life around that conviction.

Digging into BCI over the past decade has brought back the **same sense of recognition**. The evidence remains scattered across laboratories and modalities, and the technology still looks crude enough to dismiss. Yet I have seen what happens when capability compounds faster than consensus, and BCI now looks to me the same way AI did in 2018: underestimated and closer than it appears.

⁹. Still to this day the best book on the topic of AI, I highly recommend reading it.

Eight-year Offset

To borrow the AI timeline for a second: BCI has roughly an 8-year delay relative to AI; it is at the same maturity level as AI was in 2017. The year the transformer architecture¹⁰ emerged, two years before GPT-2 made scaling laws visible and five years before ChatGPT made it obvious to everyone else.



The parallel runs deeper than timing, **the brain and AI have had a deeply intertwined history since their very inception**. AI was initially *modelled on ourselves*, or on our brain to be more accurate; neurons, neural networks, synaptic weights, and training are all concepts borrowed from biology. AI took inspiration from that architecture to build its foundations. However, this flow is about to reverse, **it is now time for BCI to learn from AI to make the next leap forward**. Specifically, to use the AI scaling paradigm for the BCI field. Most people, including a meaningful chunk of the neurotechnology field itself, still do not recognize how impactful it would be to apply scaling to BCI. But scaling is about to become the dominant narrative in BCI as well.

¹⁰. The Transformer is the neural-network architecture introduced in 2017 by Google in the [Attention Is All You Need](#) paper. It replaced the recurrent and convolutional networks that came before with a simpler mechanism called *attention*, which is the precursor of every modern large model.

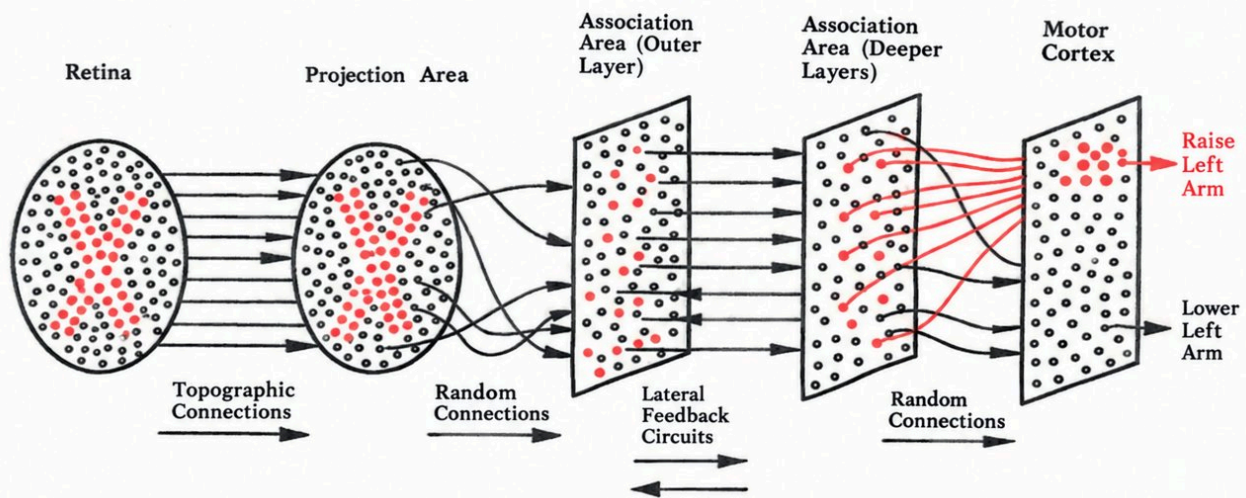


FIG. 1 — Organization of a biological brain. (Red areas indicate active cells, responding to the letter X.)

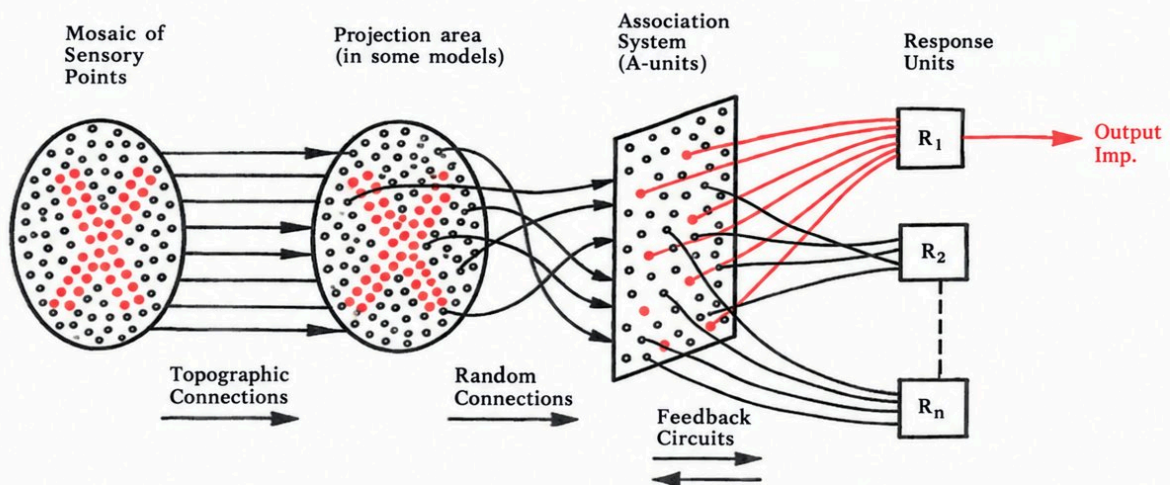


FIG. 2 — Organization of a perceptron.

A biological neuron alongside an artificial one, the resemblance is hard to miss.

The Ignition Stack

The running joke in the field is that “BCI has been five years away for 30 years”,¹¹ but the joke may finally be running out of material. The conditions for scaling BCI have been converging in the past few years, specifically:

¹¹. Neuroscientists are not the only optimists in the room, the same affliction has plagued nuclear fusion, which has been “ten years away” since the 1950s.

- **AI models got good enough.** Brain decoders are only as strong as the foundation models behind them. Until roughly 2024 those models were too small or too narrow to pull structure from biological noise. Today's frontier-class models are finally able to.
- **Sensors got cheap.** EEG, fNIRS, sEMG, the first consumer-grade OPM and ultrasound front-ends have all dropped significantly, riding smartphone and wearables curves.
- **Brain decoding graduated from toy technology to scalable technology.** The previous generation of work topped out at simple motor commands: cursor left, cursor right, one character at a time. The new generation reads sentences, images, and speech directly from brain activity. [Tang](#) ▶, [Willett](#) ▶, [Metzger](#) ▶, [Scotti](#) ▶, and the [ENIGMA](#) ▶ family each demonstrate this from a different angle.
- **Capital is finally available.** Neurotech venture funding hit \$2.3 billion in 2024 and ~\$4.8 billion in 2025. Cumulative private capital now exceeds cumulative US federal BCI-specific funding. VCs, frontier labs and governments have all been ramping up their investments in neurotechnology.

Although BCI has existed at the margins of science for decades, it is only now approaching an inflection point. Thanks to advances in sensors, foundation models, and large-scale neural data, it is beginning to be possible to do what once only belonged to speculative fiction: human thought can become a native interface to machines, and eventually the extension of the biological brain through an external cognitive layer, or *exocortex*¹².

CHECKPOINT I

[Part I](#) ▶ focused on presenting the structural realities of AI and BCI. Specifically, we have covered the following ground:

1. **The Existence Proof.** BCI is feasible. Modern computers can already decode raw meaning and visual images from the brain's biological noise; the physics allows it.
2. **The Bottleneck.** Human biology is slowing down relative to AI. As machine capability scales vertically, expanding the human-machine bandwidth and upgrading the biological cognitive system is the only way to bridge the gap and remain relevant.

12. An exocortex is an external digital layer of your mind that extends your memory, attention, and mental computation beyond your biological brain while functioning as part of it. Your phone has served as a crude, indirect version of this for over a decade; a BCI-linked AI could become a faster, more integrated version of it.

3. **The Biological Bar:** To become a consumer reality, non-invasive BCI has to clear a major hurdle: it must first beat the speed and accuracy of the peripheral nervous system.
4. **The Hardware.** The brain can be read, but surgery does not scale. Non-invasive BCI hardware is the only viable short-term path to scale.
5. **The Eight-Year Offset.** BCI is exactly where AI was in 2017. The foundational proofs of concept are starting to work in labs, but the world has not noticed yet.
6. **The Timing.** The compute is available, the AI model exists, the sensors are cheap, frontier AI labs are hitting the text-data wall and regulations are favorable. The conditions for ignition are fully set.

At this point, the theoretical debates are mostly over; the physics allows non-invasive brain reading. What remains is an execution problem; to cross the consumer utility threshold within the next five years, the BCI field must fundamentally change its architecture. It has to stop being a bespoke medical science and start acting like a scale engineering problem.

It is now time to unleash the AI scaling playbook on the human brain.



PART II.

The Scaling

Does the AI scaling playbook work on the brain?

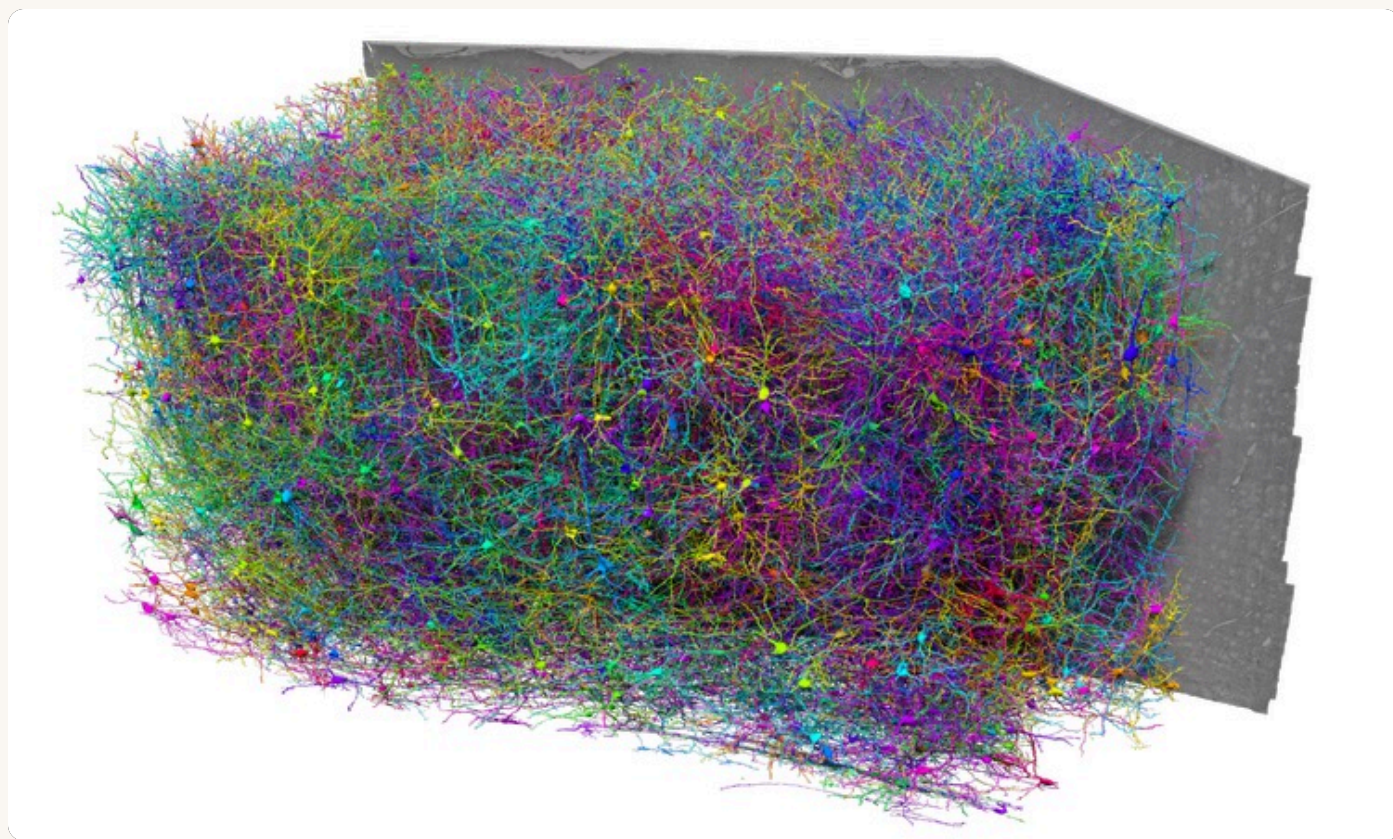
06 The Brain's Bitter Lesson

“The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin.”

— Rich Sutton, The Bitter Lesson, 2019

Let's strip away the biology for a second. If you look at the brain purely as a computer scientist, what is it?

It is a massive graph of densely connected nodes (neurons) firing off noisy electrical signals across multiple time scales. It has extreme redundancy and almost no labels. To decode the brain, you have to extract meaning from a giant, chaotic graph of raw data.



One cubic millimeter of cortex with every neuron individually traced.

But we have seen this exact architecture before; we even already solved it.

We just did not call it a brain. **We called it the internet.**

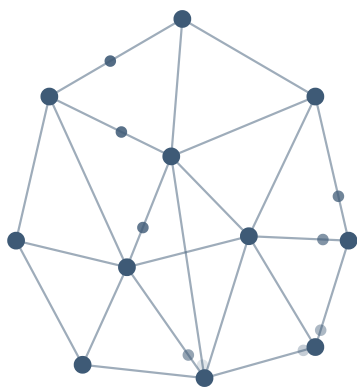
The internet is a vast ocean of unlabeled text, images, connections and noise, far too big for human engineers to manually sort. But we made it tractable, and true to Sutton's rule, we didn't do it through clever, hand-crafted tricks.

We did it by scaling an embarrassingly simple technique relentlessly: pointing very large models at very large amounts of data, and letting them figure it out.

THE GRAPH PARALLEL

The internet

unlabelled, messy, huge



The brain

unlabelled, messy, huge



$\approx$

Billions of noisy, densely connected, unlabeled nodes. The substrate is different, but the math to decode it is exactly the same.

The brain is the exact same shape of problem. It is a graph. Messier than the internet, denser than the internet, and profoundly harder to instrument. But the fundamental math is the same. Which means the playbook that conquered the silicon graph should work for the biological one.

This is no longer just a compelling handwavy theory, **it is a published, measurable reality.**

In early 2025, researchers at Meta's AI lab released a paper titled [Scaling laws for decoding images from brain activity](#)¹³. They fed an AI decoder nearly 500 hours of brain recordings, comprising 2.3 million trials across EEG, MEG, and fMRI scanners.

What they found is AI's Bitter Lesson¹⁴ playing out in biological real time:

- **The Straight Line.** Decoding accuracy scales log-linearly with data, with **no plateau in sight**. Just like a language model, if you feed the decoder more data, it gets predictably better at reading the

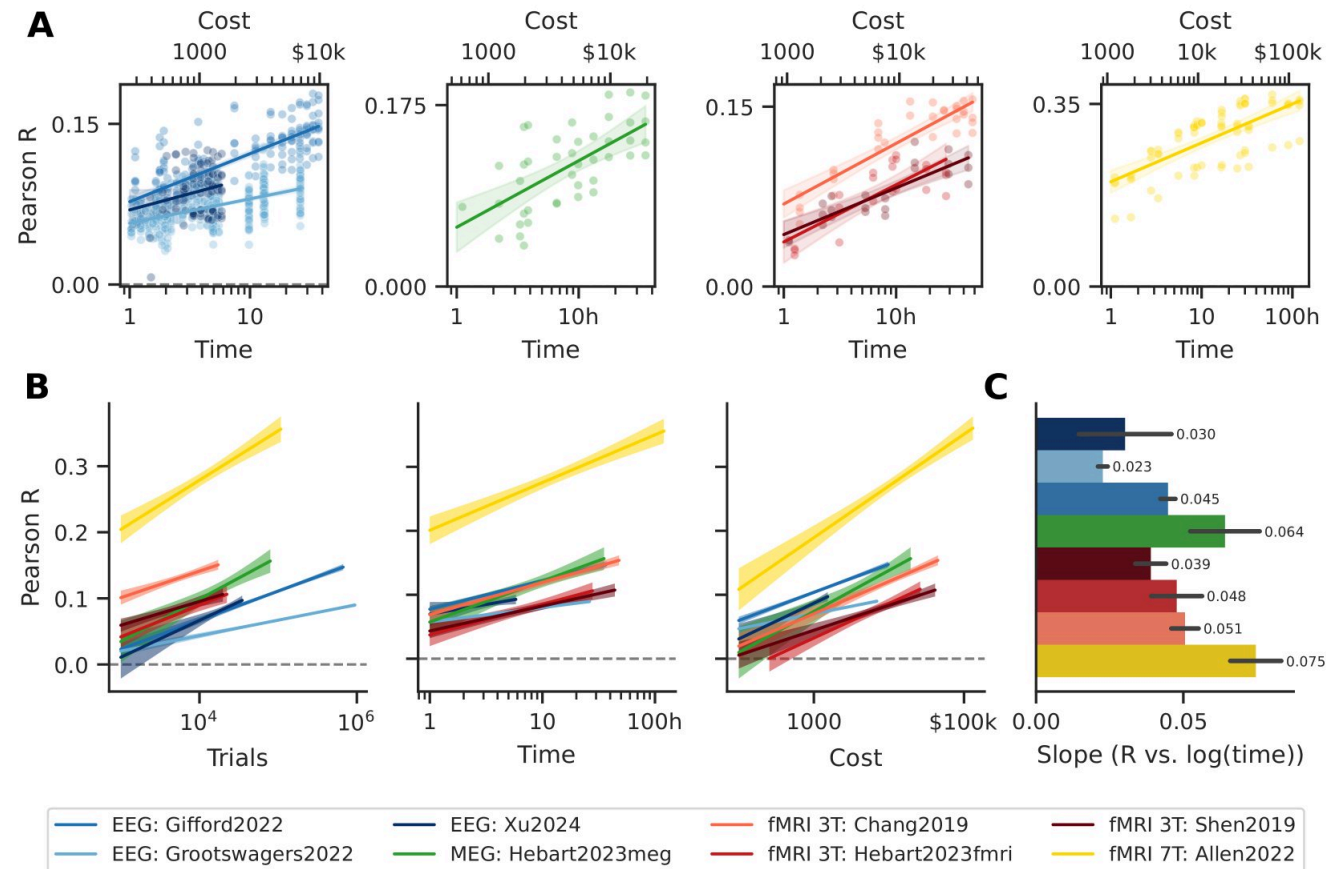
¹³. Despite its recent missteps in the metaverse and AI, Meta has been impressively ahead of the curve in neurotechnology, as its [latest research](#) demonstrates.

¹⁴. [The Bitter Lesson](#), articulated by Rich Sutton in 2019, is that general methods that scale with data and computation eventually outperform handcrafted human knowledge. It is "bitter" because researchers repeatedly resist the discovery that their clever domain expertise matters less than brute-force scale.

mind.

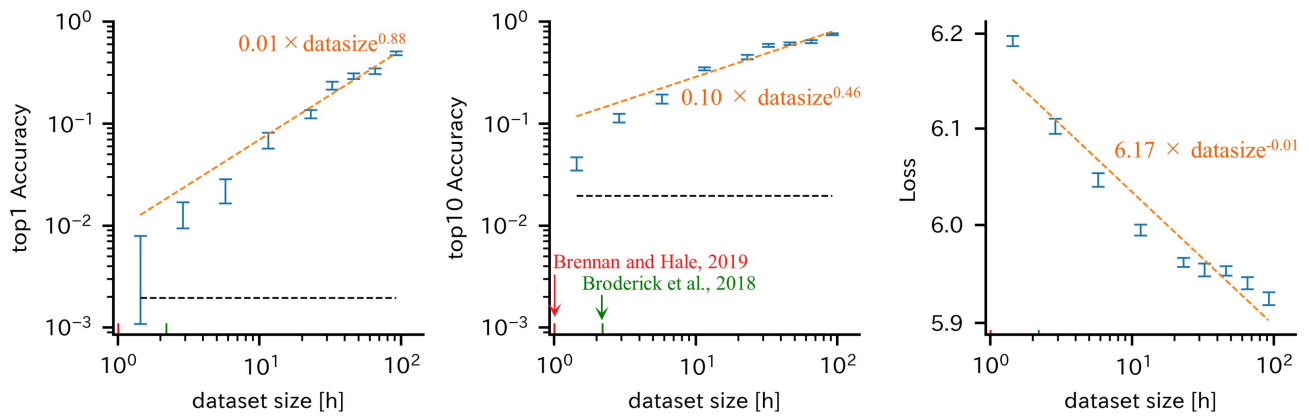
- **The Hardware Offset.** Higher-fidelity sensors (like an fMRI) simply raise the line compared to lower-fidelity sensors (like an EEG). The scaling law holds regardless of the sensor.

HOLDS ACROSS MODALITIES



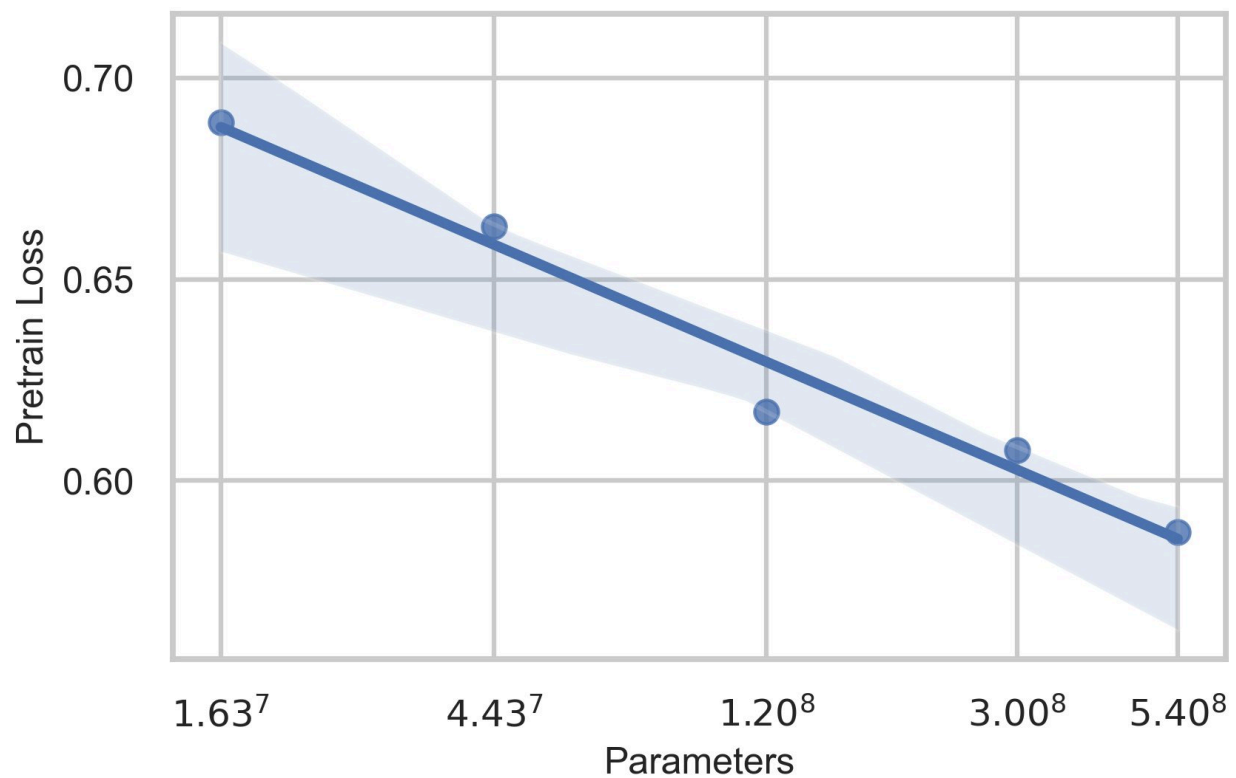
The same upward trend across ~8 EEG, MEG and fMRI studies. Source: [Banville et al., 2025](#) ▶.

POWER-LAW FIT



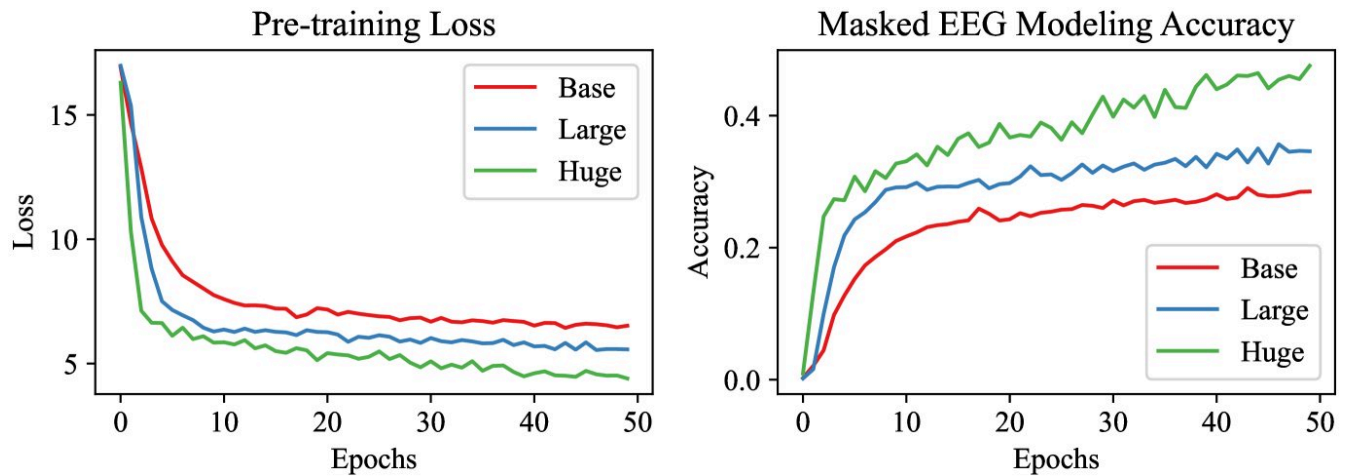
Decoding accuracy keeps climbing as you add hours of data. Source: Araya et al., 2024.

BIGGER MODELS, LOWER LOSS



Bigger models improve steadily and predictably. Source: ALFEE, 2025.

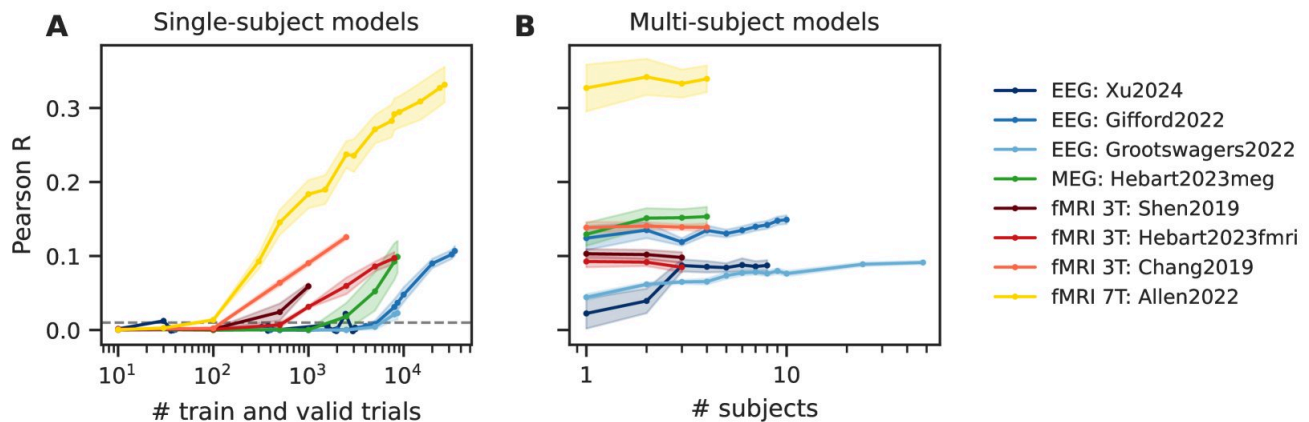
PRE-TRAIN LOSS FALLS



Bigger brain models learn more from the same EEG. Source: LaBraM (Jiang et al., 2024).

- The Data Shape.** At fixed total hours, recording one subject for 100 hours drastically outperforms recording 100 subjects for one hour. This is expected: below a certain amount of data per brain, a model cannot yet generalize across people, much like at a small scale a language model would learn far more if you trained it on 100 books in one language rather than one book from 100 languages. With scale, this constraint will disappear.

DEPTH OVER BREADTH



Below a certain amount of data per brain, a model cannot yet generalize across people. Therefore, with a fixed training budget, going incredibly deep on one single brain is much better than spreading that data across many. Source: [Banville et al., 2025](#).

That is the structural reality that will shape the future of the entire field. The scaling laws do not care about scattered, one-off clinical trials; they demand deep, continuous data. What we currently need

for decoding the human mind is not a multimillion-dollar hospital scanner collecting sparse but highly accurate data, what we need is **a cheap wearable BCI that can collect data at scale.**

We should not wait 70 years to learn the lesson that AI had to learn the hard way. We can skip the hand-crafted era of BCI entirely and jump straight to the scaling playbook.

Part II ▶ of this essay will focus on exploring how the AI scaling strategy can be applied to non-invasive BCI in order to create a first generation of consumer products and set a powerful [flywheel effect](#) ▶ in motion. Over the next 5 to 10 years, this flywheel effect will in turn accelerate the momentum of the field significantly and help BCI close the gap with AI.

Whiteboard Skepticism

Of course, this is not a guaranteed victory. Brain data is noisier and more variable across subjects than text. It is time-correlated across multiple scales, and the labels are much weaker than simply predicting the next token. Yes, when scaling laws have been tested at a small scale in BCI (through datasets like ZuCo, DEAP, and AllJoined), they appear to hold, but small-scale trends do not automatically extrapolate.

On top of it, right now, there is no such thing as a decoder that works on a new brain out of the box. Every model built so far still requires a calibration phase specifically tuned to your brain before it actually works. The field has never had to merge tens of thousands of unique human brains into one universal training set without relying on per-subject calibration.

SIDENOTE

The most counterintuitive aspect of our current ability to decode the mind is that we don't even fully know how the brain computes. We built the internet, so we understand how it works but we do not know for sure if cognition is reducible to simple neuron electrical impulses, or if sub-neuronal physics actually matters. Microtubules, glial signaling, and even quantum effects are all still live hypotheses as underlying mechanisms for thoughts. All we know is that if we put a sensor on a head, grab a bunch of messy data, and ask an AI to find the pattern, it actually starts reading minds.

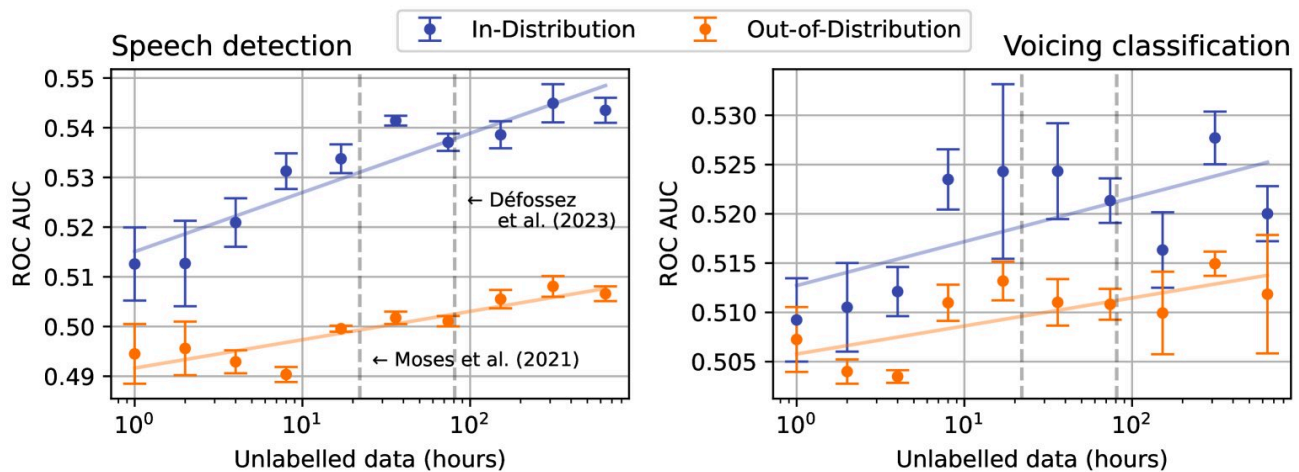
The scaling playbook is a high-stakes wager, it is a bet that brute-force compute can extract enough usable structure from the noisy macro-signals of the brain without needing a perfect map of human cognition first. **A bet that we can decode the mind without fully understanding it.**¹⁵ If this turns out

to be right, the scaling lines will continue pointing upwards and new breakthroughs will be achieved in fast succession as data and compute scale.

If the last decade of AI taught us anything, it is that betting against compute makes you sound incredibly smart, *right up until it does not*.

Theoretical counter-arguments can only go so far here. All the caveats above simply tell us that the brain is not exactly the internet, which is already obvious from the premise. The thesis is that it is *similar enough* that the exact same laws of compute apply. **But you fundamentally cannot disprove scaling laws with a whiteboard.** It is inherently an empirical problem. The only way to prove that scaling fails is to actually run the experiment and hit a wall.

SCALING TO NOVEL BRAINS

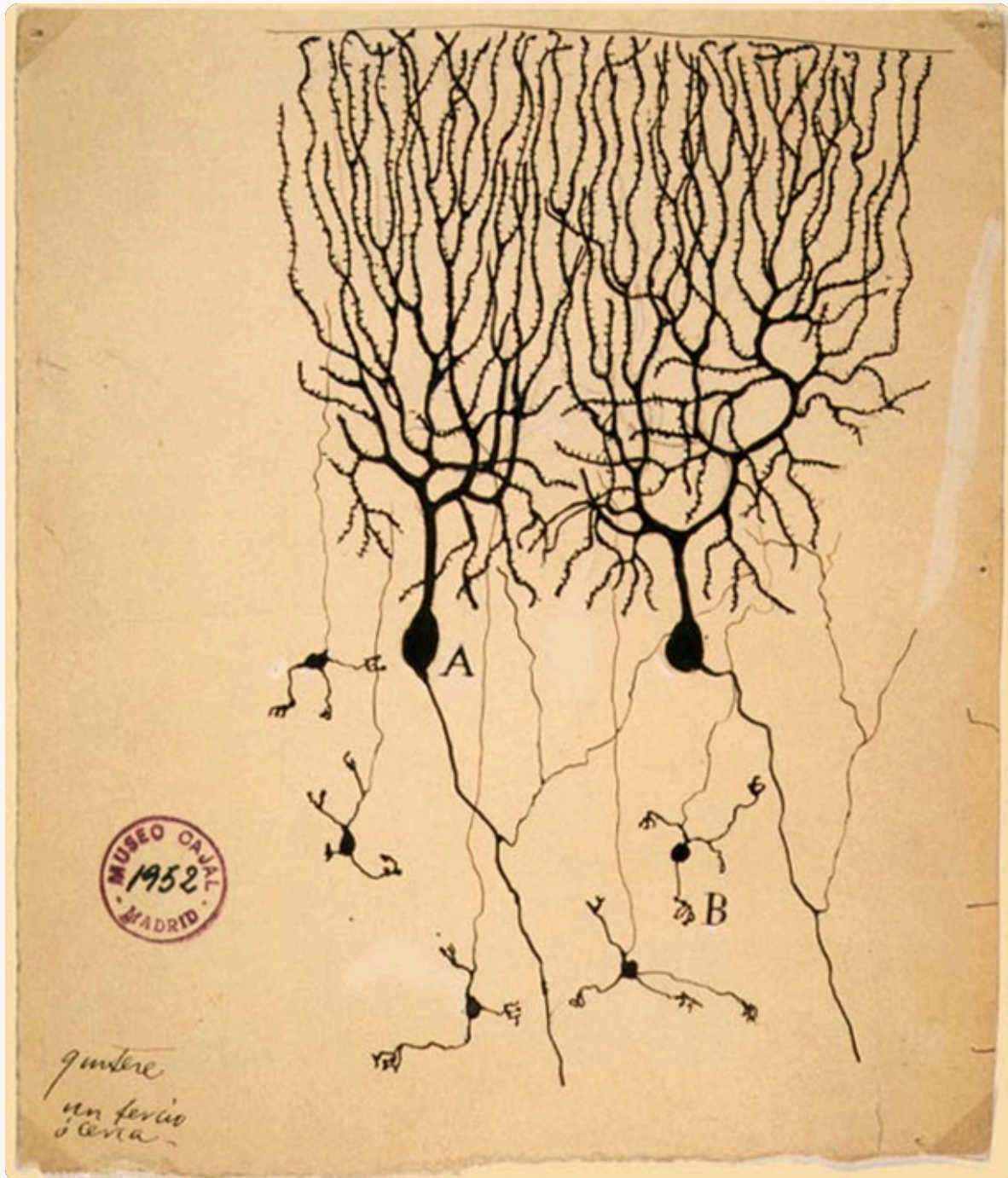


Independent replication, Oxford researchers ran the hardest possible version of the scaling test by evaluating the model on completely novel brains it had never seen before and the straight line still held. The researchers were so convinced by the math that they named their paper [The Brain's Bitter Lesson](#) ▶.

And so far, every attempt to scale has only reinforced the evidence.

[The Bitter Lesson](#) ▶, as Sutton put it (or its older sibling, Peter Norvig's [The Unreasonable Effectiveness of Data](#) ▶, which made the exact same point in 2009), will come for every field eventually. It has not yet come for BCI; it is overdue.

15. There is a deeper question lurking here: whether “understanding the brain” is even a well-defined milestone. I believe it is not: understanding is compression, and there is only so far a system as complex as the brain will compress. Tal Yarkoni makes the fuller case, and argues we might not even notice reaching it, in [If we already understood the brain, would we even know it?](#)



The first known schematic of the brain. Santiago Ramón y Cajal, 1899.

Why no one has run the playbook yet

If the scaling bet is so obvious, a glaring question remains. BCI has had decades, why has nobody actually run the playbook?

There are a couple of reasons.

First and foremost, the specific AI architectures and self-supervised techniques required to actually execute this strategy were only proven in the last few years.

But this is not the only reason, this stopped being a blocker a few years ago. What explains the delay in adopting this strategy is institutional inertia. Culture evolves at a glacial pace, or as Max Planck colorfully puts it: *“Science advances one funeral at a time”*.

The entire BCI field was born in the world of invasive medical devices. Going from a culture shaped by FDA approvals, extreme patient safety and statistical rigor on small datasets to one of brute-force and scale is a **brutal paradigm shift**. Researchers who spent their careers learning to carefully hand-craft algorithms for specific patients do not naturally reach for the “throw 100K hours at a wall of GPUs” strategy. It certainly has a lot less scientific appeal. This heavy medical baggage was built for brain surgery, but it has completely bled over into the non-invasive space, and it is a real drag on the field.

Furthermore, a lot of the hardware for non-invasive BCI only recently became cheap and powerful enough for this purpose. Historically, the highest-performance, most “serious frontier” BCI results have been invasive or semi-invasive. Which explains why non-invasive hardware took a while to get the attention it deserved. There has also been a lab-fit hardware culture. Research-grade hardware optimizes for signal quality in a controlled setting; scaling-grade hardware optimizes for daily wear at the expense of signal quality.

Lastly, large coordinated data collection efforts only started to come to life recently. There is no Common Crawl¹⁶ or ImageNet¹⁷ for the human brain. We have several partial equivalents, but they are fragmented, small, clinical, lab-specific, and rarely track the same person over a long period of time. Data fragmentation is a major issue. Each lab collects its own dataset with its own preprocessing, electrode layouts, and approval pipelines. Some initiatives like BIDS and NWB are nascent but adoption is uneven. A paper saying it trained on “eight datasets” usually means a poor graduate student spent six miserable weeks reformatting files before training even began.

A lot of these are coordination problems. Coordination problems get fixed the moment one or two actors finally make the new optimal visible, forcing the rest of the field to follow, just like OpenAI and Google did for AI. The reason the scaling playbook has not run yet is not because it violates the laws of physics. **It is because no coordinated push has been done in this direction**, but that will change this decade.

¹⁶. Common Crawl is a non-profit that has been continuously crawling and publishing the web since 2008; the resulting petabyte-scale dataset is what most large language models trained on.

¹⁷. ImageNet is a labeled dataset of ~14 million images across 20K+ categories. The 2012 ImageNet competition is widely credited as the beginning of the deep-learning era in computer vision.

The best strategy for the BCI field is not to try to be clever about brain decoding, it is to be willing to scale.

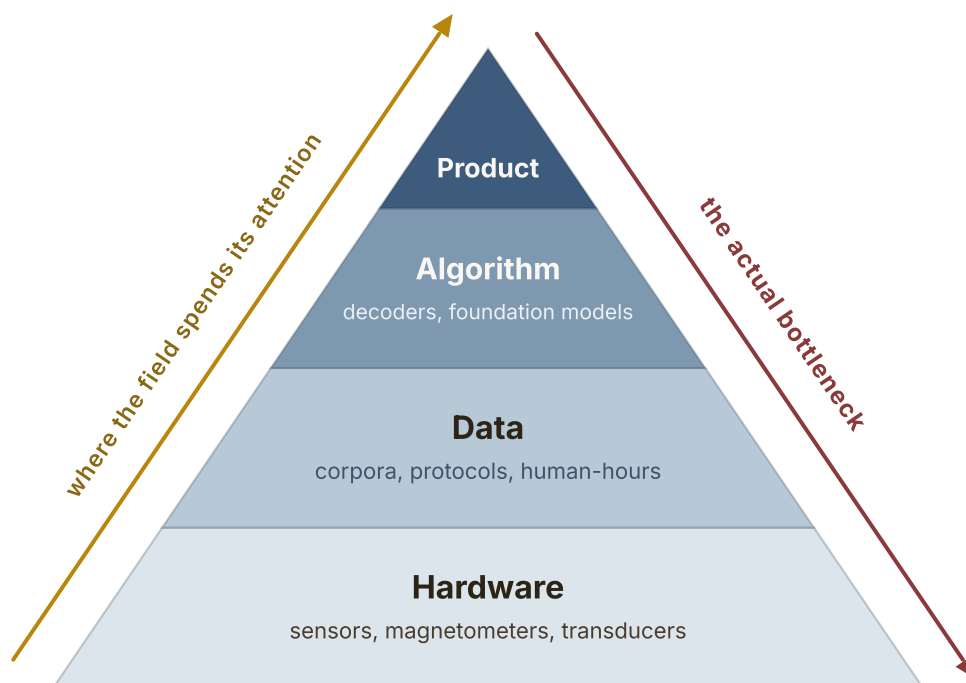
The natural next question is: how do we actually apply scaling in practice. To map the execution out, we are going to use a simple mental model, the BCI pyramid.

07 The Pyramid

Before diving into the details of how to scale non-invasive BCI, we need to break it down to its core. Fundamentally, *what is the field of non-invasive BCI made of?*

The simplest mental model is a pyramid. It has four layers, with every layer built directly upon the foundation of the one below it: Hardware, Data, Algorithms, and Product.

THE BCI PYRAMID



phaseLock.satpugnet.com

Each layer is bottlenecked by the one below it. Most of the field's attention has been at the top of the pyramid.

Hardware is the base. Sensors, magnetometers, transducers, hardware sets the ceiling on signal quality. The rule is simple: anything you cannot record, you cannot decode. Right now, this layer is severely underserved. The equipment is often bespoke, expensive, and hacky. **The foundation of the entire field is still very underdeveloped.**

Data sits directly above that. The corpora of recordings, the protocols, the formatting conventions. Data sets the ceiling on how smart your model can get. You can only perform so much mathematical magic with 100 hours of recordings. Just like the hardware beneath it, **this layer is critically immature and fragmented.**

Algorithm sits above the data. These are the AI models that turn raw signal into actual meaning. Better algorithms can compensate for noisy data, but only up to a point. Today, a massive amount of the field's time and talent is wasted here, trying to squeeze every last drop of signal out of tiny datasets.

Product sits at the very top. It is entirely constrained by the three layers below it. Most of the products built in non-invasive BCI today have severely limited capabilities because the foundations they sit on are not robust. **Building a consumer BCI right now is like trying to build ChatGPT when you only own a hundred GPUs and ten gigabytes of text.** It is a structurally impossible task. The companies doing it are working miracles, but they are fighting a brutal systemic headwind.

The core dynamic here is simple, **each layer is bottlenecked by the one below it.** You cannot ship a product that requires data your hardware cannot collect. You cannot train a foundation-grade decoder on a couple hours of a graduate student staring at a screen in a basement.

The problem with the field today is that **too much attention is at the top of the pyramid and not enough at the bottom.** Everyone is obsessing over the penthouse while the foundation is sinking. The hardware layer is starved of well-funded and focused companies. The data layer needs aggressive scaling. And yet, every month brings dozens of beautiful papers proposing new decoders trained on tiny datasets. The algorithm is not the main problem. **Data is the problem, and the hardware that collects the data is the problem before that.**

Take the [ZuCo dataset](#) ¹⁸, every time a research team works a mathematical *miracle* to squeeze a new paper out of the brain data of 12 people reading for six hours¹⁸, they are demonstrating the field's problem more than they are solving it. Squeezing blood from a stone is impressive, but at a certain point, the miracle just becomes proof of the bottleneck. **We need to stop praising the squeezing and start asking why we only have stones.**

Now that we have a clear mental model of the field, we are going to dissect exactly what is causing the bottleneck at the top. We will *walk up the pyramid layer by layer*. First the hardware, then the data, then

¹⁸. Yes, I promise you this is not a joke. One of the most commonly used datasets in the entire field literally consists of 12 people recorded for six hours each. That is 72 total hours of data. We are trying to build the ultimate foundation model for the human brain using the equivalent of one long weekend of recordings. It is like teaching an AI Japanese by showing it one sushi menu.

the AI algorithms, and finally, the complex coordination required across the entire pyramid to actually pull this off.

08 Level One: Hardware Is Hard

“People who are really serious about software should make their own hardware.” — Alan Kay, 1982

Let us begin at the bottom of the pyramid: **Hardware**.

AI hardware has been all over the news recently. We watch trillions of dollars flow into massive GPU data centers because everyone now understands that compute is the ultimate bottleneck for artificial intelligence. But hardware is not just the blocker for AI, it is also one of the key gating factors for brain-computer interfaces, and right now, the BCI field desperately needs its *own hardware paradigm shift*.

A device engineered to gather massive datasets is a fundamentally different species of machine than a device engineered for a laboratory, and one of the most important differences is portability. If you are optimizing for the *absolute cleanest signal possible*, you demand precision at all costs. You accept the electromagnetically shielded room, the liquid helium cryogenics, and the completely immobile subject. However, if you are optimizing for the largest dataset possible, the math changes, you still care about signal quality, but the trade-off tilts heavily toward **portability**. Slightly noisier signal collected at a thousand times the volume beats a clean signal you cannot collect outside a shielded room. That is the trade-off the field has not yet fully internalized.

A slightly noisier signal collected at a thousand times the volume completely beats a pristine signal you cannot collect outside a shielded room

Today, the current state of the art is *heavy, expensive*, and lives in a *room with no windows*, and not just metaphorically. That is fine for a paper, but it is incompatible with collecting tens of thousands of hours per subject, across thousands of subjects, in real-world conditions. The field needs a person to be able to put a helmet on in the morning, go about their day, and come home with eight hours of multi-modal neural data perfectly paired with the ambient video and audio.

This does not necessarily imply that all data should be collected in fully uncontrolled environments with hardware that could withstand anything. Reality has a long tail of edge cases. In practice, in the

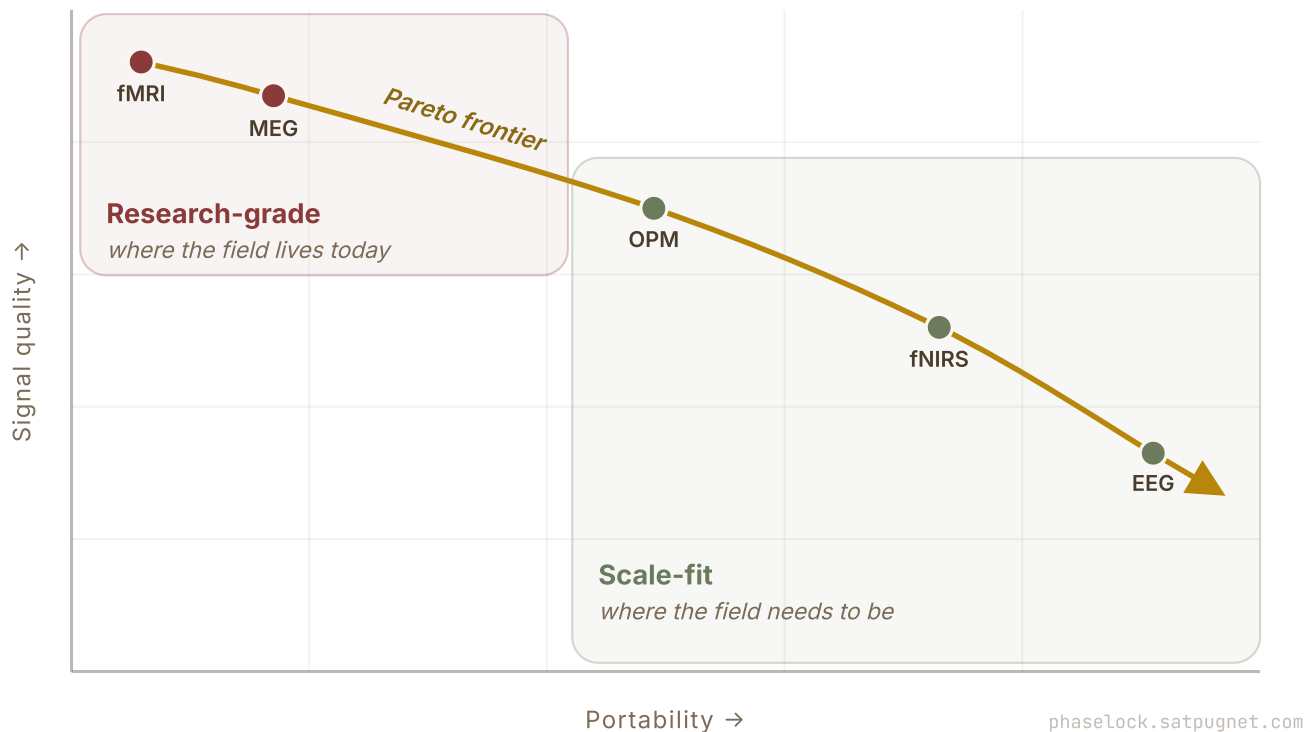
short term, a significant amount of data will still be collected in semi-controlled environments. But even with marginal improvements, more portable hardware unlocks the flexibility to seamlessly record across a wide range of scenarios, that **flexibility** is exactly what can **compound into vast, robust datasets**.

Pristine lab data will still matter, doing model fine-tuning and evaluation with high quality, low noise data still provides a significant advantage¹⁹, and we will always need those clinical devices. But for the vast bulk of the training, having **slightly noisier data at massive scale is much more valuable**.

SIDENOTE

Of course, there is a limit to how much noise a dataset can have while remaining useful, and historically, consumer neural recordings have been *10x or even 100x noisier* than their lab counterparts, reducing their practical usefulness for training. However, as hardware is now improving quickly, that limitation is disappearing fast.

THE SCALE-FIT TRADE-OFF



The Pareto frontier of signal quality vs portability; the optimal is the bottom right, not the top left.

Two things follow.

¹⁹. The best strategy for training large brain models probably mirrors the one used in autonomous driving: train the base model on abundant but imperfect data (in their case simulated data), then fine-tune and validate it on smaller volumes of high-quality data (in their case real-world data). For BCIs, that would translate to pretraining on noisy consumer recordings, then fine-tuning and evaluating on cleaner lab data; [there is already evidence](#) that this works.

First, the field needs more specialized hardware companies. Companies that focus on the scaling paradigm rather than the highest-signal-quality-at-any-cost paradigm. Hardware iteration is its own brutal discipline, treating it as a side-effect of an algorithm paper has held the field back for two decades. Conway's Law²⁰ applies here: organizations ship their structure, and an academic field has shipped academic hardware. Thankfully, more specialized hardware companies are finally seeing the light of day, the momentum is already shifting away from signal purity toward scaling.

Second, the right design choices are different when scale is the goal. If portability is the first pillar of scale-fit hardware, **multimodality is the second.**

Many modalities can be used in tandem. MEG, fNIRS, and EEG measure entirely different physical phenomena; they do not interfere with each other, so you can stack them. When you are aiming for massive, robust datasets, capturing multiple biological signals at a low marginal cost is an obvious choice. The same way that modern LLMs are trained on text, videos and audio, *brain foundation models should learn from multiple complementary biological signals*. Multimodality should not be just a clever trick, it should be the sensible default.

SIDENOTE

Additionally, in the past, models and sensors were mostly developed *independently*, with people making educated guesses about which changes to noninvasive sensors would most improve them. But if we now train large models directly on the sensor's output, we don't have to guess anymore, we can **co-design sensors** and models to create a very *tight feedback loop*.

Once you accept portability and multimodality as your foundation, the rest of the engineering constraints follow. Building for scale means solving for:

- **Manufacturing.** How do you actually ship a million units?
- **Context.** Embedding cameras and microphones to give the neural data real-world grounding.
- **Tolerance.** Ensuring your sensor array can handle hair, sweat, physical motion, and the chaos of daily life.
- **Standardization.** One of the most critical and underweighted components. If your hardware outputs data that cannot be cleanly merged with another lab's recordings, the data layer above it stays hopelessly uncoordinated.

The hardware bottleneck is slowly being overcome, but unevenly.

Each of the major non-invasive modalities (EEG, ultrasound, fNIRS, OPM-MEG, eventually NV-magnetometry) has started making the jump from "lab" to "wearable". [Butterfly Network](#) put ultrasound on a chip, [Kernel](#) built a wearable fNIRS rig, [Cerca Magnetix](#) ships OPM arrays you wear.

20. Conway's Law: software systems mirror the communication structures of the organizations that build them. Coined by Melvin Conway in 1968.

THE MODALITY MAP

MODALITY	SPATIAL	BANDWIDTH	TEMPORAL	PORTABLE	COST	DEPTH	MATURITY
● fMRI							
● Intracortical							
● MEG / OPM							
● Ultrasound							
● ECoG							
● sEEG							
● NV-center							
● Endovascular							
● fNIRS							
● EEG							

phaselock.satpugnet.com

No modality dominates: EEG leads on cost and portability, fMRI on spatial resolution, and ECoG and intracortical on signal quality. The rest offer different compromises across depth, speed, and maturity. [Appendix D](#) describes each one in full.

Hardware is also where the geopolitical reality sets in. China [has been investing aggressively](#) in BCI hardware over the last several years. Their [five-year plan](#) put BCI squarely inside the national strategic technology stack. China does not see BCI as a neuroscience curiosity anymore, it already sees it as a strategic asset.

We have seen this movie before, the AI decade was a fight over chips; the defining geopolitical fact has been who can make them; this has come down to one Taiwanese fab, one Dutch lithography company, a handful of US export controls, and a trillion dollars of strategic anxiety on top. This race has already started for BCI hardware, if brain data becomes strategically important, the machines that collect it will become strategically important too. **Within a few years, the TSMCs of brain hardware will start taking form**, and China is clearly trying to make sure the answer is not automatically American.

To sum it up, the point of this section is that the hardware unblock matters more than another decoder paper. The next leap in non-invasive BCI will not come from a smarter algorithm trained on a tiny dataset. It will come from sensors that can *collect data at scale*, sensors which can be *worn by many people, for extended periods of time*.

But sensors alone only produce raw numbers. Once the hardware to capture the signal at scale is no longer the bottleneck, the next one becomes: how to collect it at scale.

09 Level Two: The Data Scaling Argument

“Simple models and a lot of data trump more elaborate models based on less data.”

— Halevy, Norvig, Pereira, The Unreasonable Effectiveness of Data, 2009

The second layer of the pyramid is: **Data**.

For decades, the AI field chased clever symbolic methods, largely because it did not realize that brute force was a real option. The Bitter Lesson highlighted that mistake. The most uncomfortable thing about this outcome is that it is a deeply unsatisfying scientific result: the elegant theory of language did not produce GPT, scaling mindlessly did. That is in part why fields like BCI take so long to internalize the move.

BCI today has a significant advantage over AI in 2017: the AI model *already* exists. It simply needs to collect the data. The field is starting to awaken to this and is currently crossing into its **foundation-model era**.²¹

DECODE-IT-YOURSELF: THE SCALING LAW, ON ONE SENTENCE

WHAT THE SUBJECT HEARD

I got up from the air mattress and pressed my face against the glass of the bedroom window.

DECODED · ~10 HOURS

he saw a dog then looked
out the window

DECODED · ~100 HOURS

I woke from bed and put
my face to the glass
window

DECODED · ~25,000 HOURS

I got up from the air
mattress and pressed my
face against the glass of
the bedroom window

■ error ■ gist ■ exact

phaselock.satpugnet.com

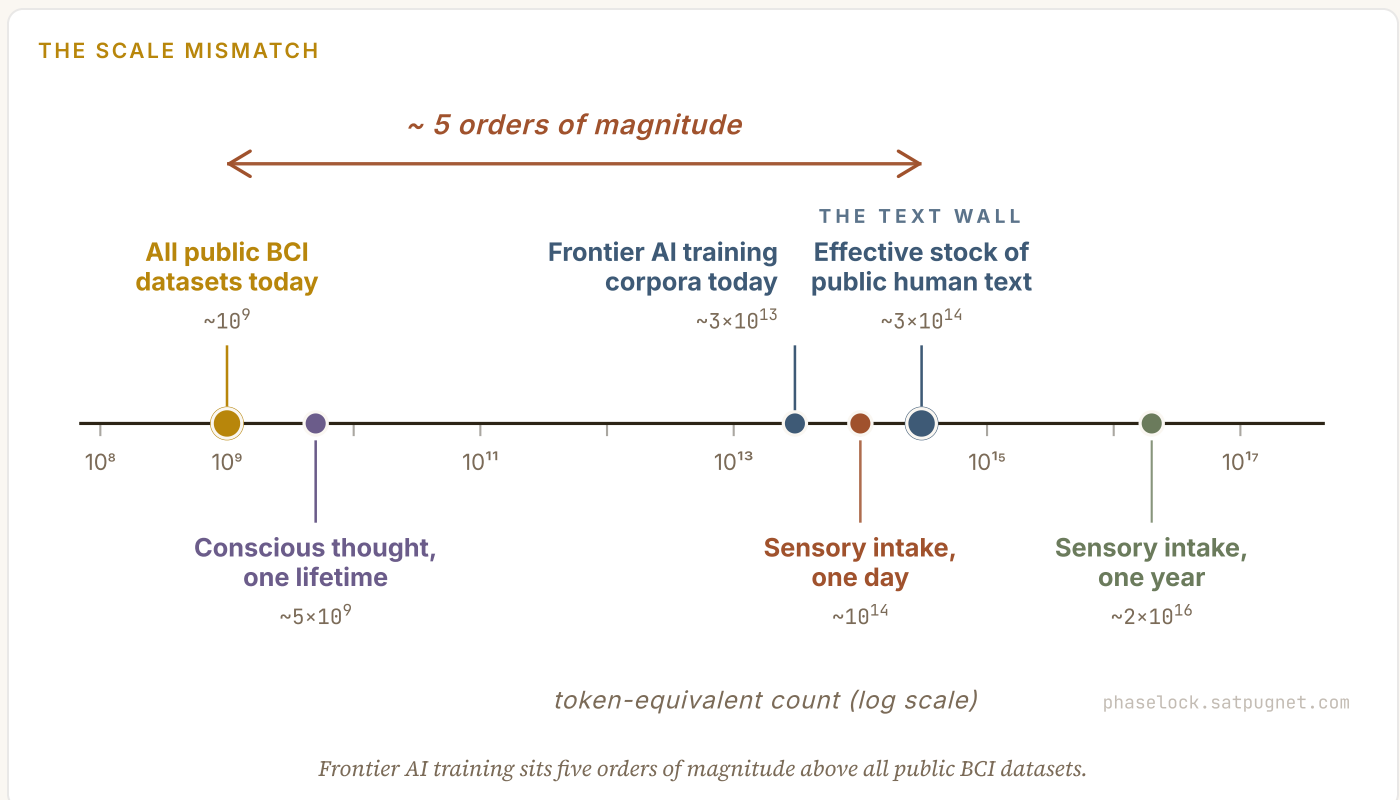
More data leads to sharper decoding.

As the hardware is improving, scaling data collection becomes easier and datasets are growing, but BCI datasets are still nowhere near the scale that earned foundation models their name. The largest

²¹ I have been circling this idea for a while, it just took my distracted self a while to actually get it on paper. Here are some receipts from 2024: [“We need a new Moore’s law for brain mapping. Full brain mapping will arrive sooner than we think.”](#) and [“Direct brain to LLM communication probably coming sooner than we think.”](#)

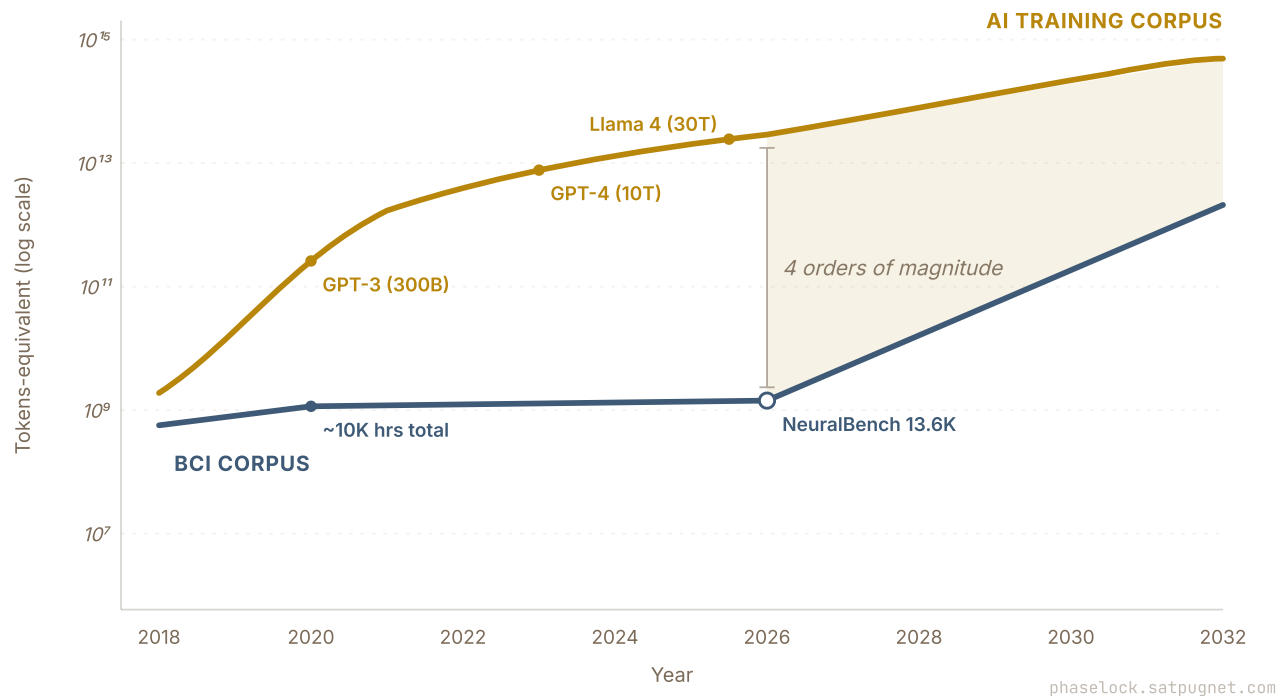
publicly available EEG datasets are tens of subjects, tens of hours each, single modality, mostly in a lab ([ZuCo](#) ▶, [DEAP](#) ▶, the [Stanford handwriting corpus](#) ▶, the [BCI Competition datasets](#) ▶). The most ambitious in-house corpora I am aware of are still *under* 10K hours total, and single modality.

However, that picture is starting to change. [NeuralBench v1.0](#) ▶ (Meta, May 2026) consolidates 13,603 hours of EEG across 9,478 subjects from 94 datasets. Models are scaling alongside: [NeuroLM](#) ([Jiang et al., ICLR 2025](#)) ▶ at 1.7B parameters on 25,000 hours of EEG; [LaBraM](#) ([Jiang et al., ICLR 2024](#)) ▶ at 369M and 2,500 hours.



For comparison: GPT-3 (2020) was trained on roughly 300 billion tokens. Llama 3 (2024) used 15 trillion; Llama 4 (April 2025), 30 trillion. The effective stock of public text caps at roughly [300 trillion tokens](#) ▶. BCI has still a long way to go to be in AI scale territory.

THE AI-BCI DATA GAP



The BCI corpora currently sit four orders of magnitude below the AI one.

If you are inside the BCI field, this data scaling argument may sound obvious. To a meaningful chunk of your colleagues, it is not. The field's culture has been *“do clever things on small datasets, publish, repeat”* for a long time. The move to *“build the boring foundation first, and scale until the Bitter Lesson kicks in”* can read as a step away from real science, it can feel less beautiful or clean, but it is also the move that worked everywhere else and that the BCI field needs.

SIDENOTE

The data collection argument compounds in one more way. The massive BCI datasets produced by following this playbook are also very valuable to AI labs. Indeed, as we get closer to the data wall and alignment becomes a higher priority, AI labs are becoming more and more interested in alternative data sources. Therefore, a large dataset of human cognitive processes is both valuable to the BCI field as well as the AI field. The thesis is not just that BCI scales, it is that *BCI scales and AI labs scale toward each other*, following a self-reinforcing cycle. We will discuss this point more in depth in [Part III](#) ▶.

As datasets grow larger, training large models to decode the brain becomes more feasible. But to actually extract meaning effectively from this scale of data, the field has to *abandon its custom algorithms* and focus on the architectures and strategies that conquered language.

10 Level Three: Foundation Models for the Brain

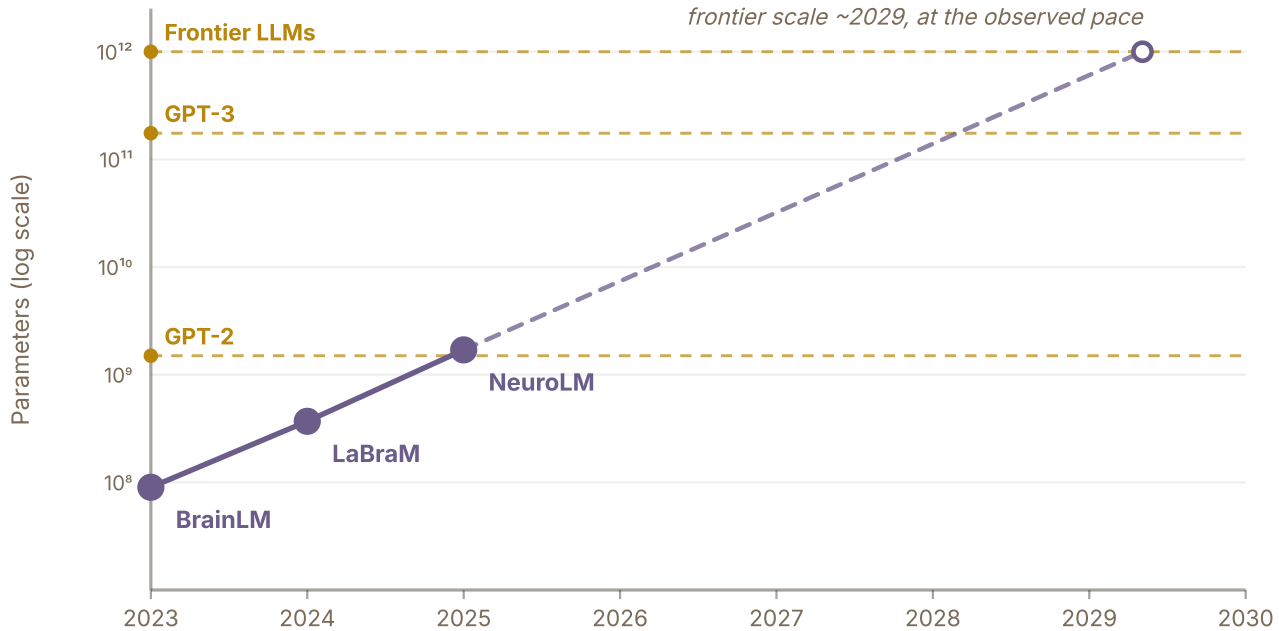
Data without models is not of much use, and so the next layer of the pyramid is the AI layer trained on the neural dataset.

The massive advantage BCI has today is that **we do not need to invent the math**. The AI playbook is already written. Little creativity is required on the architecture side, you do not need to reinvent the wheel, you take the exact transformer models that mastered language, tweak the ingestion layer to accept neural data, and hit run.²²

Another page we must steal from the AI playbook is to **build the foundation in the open**. Right now, the BCI field lacks the data to even test the upper limits of a non-invasive brain decoder through scaling. But trying to hit that scale in a silo would be unwise, the way the AI field tackled this situation in its early days is through open collaboration. The foundational work at OpenAI, DeepMind, and FAIR was strikingly collaborative. The datasets were shared, the papers were openly published, and the models were generally released. The AI field eventually moved toward closed weights, but only after the foundation was laid. BCI is still firmly in the foundations-laying phase. **Starting closed is bound to slow progress.**

22. It will obviously be more complex than this but the point is that no breakthrough is required, mostly engineering.

MODEL SIZE: BCI VS. AI



phaseLock.satpugnet.com

Brain foundation models have closed the [GPT-2 parameter gap](#). However, the data they train on is still ~6 orders of magnitude below frontier-text token volumes, which is now the harder bottleneck.

A small set of startups and academics have begun pushing in this direction²³, the *ecosystem is growing*. Because the AI field is already so mature, the models themselves are not the blocker right now. What it lacks is the massive corpus to train them at frontier scale and the empirical proof that frontier scale produces the same capability uplift for brains that it did for text.

There is also a compute tailwind that benefits BCI specifically. GPU performance per dollar has been [roughly doubling every two years](#), and the cost of training any given model capability has collapsed by orders of magnitude in five years. A GPT-3-equivalent run that cost millions in 2020 now sits in the low six figures and **keeps falling**. BCI does not need to fund frontier-AI-scale compute today. By the time a brain corpus justifies frontier compute, the compute will cost a fraction of what frontier labs pay now. The hard problem of this decade is *the data side*; the compute side gets cheaper simply by waiting.

We are now one [Common Crawl](#) and one foundation model away from BCI's ChatGPT moment.

However, the biggest challenge in running the strategy that I have described is not the individual layers of the pyramid, it is the coordination across the field. As is often the case with deep tech, the

²³. The broader open ecosystem has produced roughly a dozen brain foundation models since 2023: BrainLM, LaBraM, NeuroLM, REVE, POYO+, NDT3, LUNA, BENDR, BIOT, CBraMod, EEG-GPT, Neuro-GPT among them.

hardest engineering problem is usually not the silicon or the math, it is **human coordination**.

11 The Coordination Conundrum

Coordinated scaling is something BCI has not yet seriously attempted. It runs through the three independent layers of the BCI pyramid (hardware, data and models). Without it, every team is solving the *same* problem from *scratch*.

To understand why coordination is so important, we can have a look at the data itself. Brain data is inherently more complex than the text data used to train modern LLMs. Text is just discrete, standardized words whereas brain data is *continuous, messy physics*; it varies wildly depending on the hardware, the environment, and the human wearing the device.

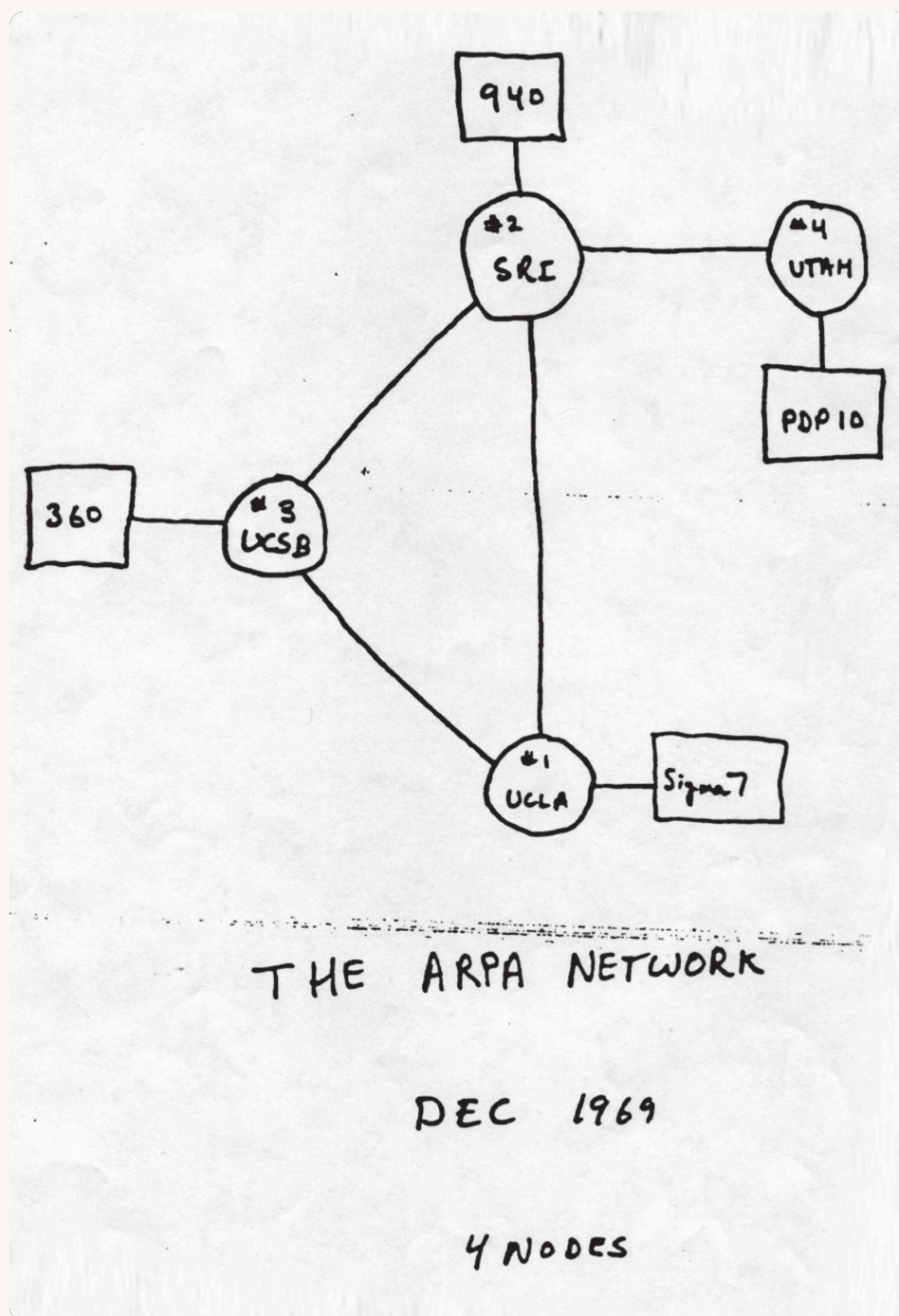
Because of this inherent complexity, most labs structure their data completely differently. There are attempts at standard formats (like [BIDS](#) (Brain Imaging Data Structure) and [NWB](#) (Neurodata Without Borders)), but their coverage is *partial*; outside their reach, formats range from MATLAB MAT files to custom HDF5 to bespoke CSV exports. The cost of this fragmentation is invisible to outsiders but painful to insiders: a paper trained on “eight datasets” usually means a poor soul spent *six grueling weeks* just reformatting files before training even began. Multiply that friction across every lab in the world and consolidating a unified dataset becomes a real challenge.

The AI data revolution did not happen by accident. **Common Crawl, ImageNet, and a handful of other artifacts standardized data collection**, and the field compounded on top of them. This meant researchers could stop wasting time collecting images and start competing on algorithms. ImageNet alone seeded a decade of computer-vision progress because everyone trained on the same images and reported the same numbers. BCI desperately needs its ImageNet. But because brain data is so much harder to standardize than text or photos (and this is where my earlier analogy comparing the brain and the internet graph bends), establishing these conventions is going to be ten times harder, and ten times more important.²⁴

Underneath the file-format problem sits a deeper structural one. The traditional neuroscience lab is highly siloed. A lab buys a few commercial sensors, runs a highly controlled experiment on 20 people, cleans the data just enough to get a result, and publishes a paper. That structure is fantastic for publishing niche scientific insights. But it is completely *useless* for building industrialized datasets of billions of hours. Academia simply is not built to be a **massive data factory**. Furthermore, the academic incentive structure still rewards researchers for building clever algorithms on tiny datasets, not for doing the boring, unglamorous work of standardizing massive ones. Expecting this system to spontaneously generate a unified, global data standard is like trying to herd academic cats. Tenured

²⁴. An early example of the kind of standardization this argument calls for: Meta AI's [NeuralBench](#), released in May 2026, is an open-source benchmarking framework that unifies 94 brain-data datasets across 14 deep-learning architectures, with initial focus on EEG and extensions to MEG and fMRI. One of the first attempts at a shared evaluation dataset for the field.

principal investigators have spent their entire careers being rewarded for being unique and clever, not for acting like a compliant, standardized cog in a massive industrial data factory. This is in some ways a *brain-coded Tragedy of the Commons*.



ARPANET in December 1969. The largest coordination system we ever built, the internet, started with just four nodes on a sheet of paper.

Coordination is not just a data-layer concern. Hardware standardization matters too: if your device's output cannot be cleanly merged with another lab's recordings, the data layer above stays

uncoordinated. Likewise, open foundation models are coordination artifacts in their own right; they let the field share the cost of training one decoder rather than every lab training their own. Each of the three layers has a coordination dimension.

And because the field is fragmented, [the three layers of the pyramid, hardware, data, and models ▶](#), are stuck in a classic **cold-start trap**.

- Hardware companies *cannot* justify scale-out until the data layer creates demand for it.
- Data collectors *cannot* gather massive volume until hardware becomes cheap and portable enough, and before BCI demand for data increases.
- Model builders *cannot* justify frontier compute costs until the massive datasets exist to train on.

Each layer is **rationally waiting** for the layer below to move first. As a result, the whole stack stays stuck at lab scale by default. The way out is for every layer to start building for the anticipated scale the layer above will **eventually require**, not the scale current demand rewards. That is the coordinated strategy transition this decade needs to unstick the cold start.

The underlying technology is finally ready. The remaining challenge is to coordinate scaling as BCI transitions from a scientific problem into an engineering problem. To succeed, non-invasive BCI now requires founder energy, massive capital, and a tolerance for *brute-force engineering* that the academic system simply was not designed to deliver.

Now, if we assume the field successfully executes this playbook, where would that actually lead us? And what exactly are we building?

12 The Two Ladders

If the playbook works and compounds, how do we actually measure success? To answer this question, we will use two distinct “ladders of progress”. The first measures what a device can actually read. The second measures how much effort a user has to invest before the device works for them.

We will call them **the capability ladder and the calibration ladder**. They represent the two main dimensions which define BCI progress.

The Capability Ladder

The capability ladder represents the raw decoding power of the BCI: how detailed the extracted information is, and how reliably the device can read it. Each rung on the capability ladder corresponds to roughly one order of magnitude better signal quality (driven by better hardware, more data, larger models), which in turn unlocks a new class of product that was previously impossible.

THE CAPABILITY LADDER



phaseLock.satpugnet.com

Each rung represents a gain in signal quality. We do not know exactly where each modality caps on this ladder, but this manifesto's argument is that the non-invasive cap is much higher than most of the field assumes.

Non-invasive hardware fundamentally has a lower ceiling on this ladder than invasive surgery. At the limit, a sensor outside the brain will never be able to beat one within the brain. A wearable sensor will never read a single, isolated neuron the way a surgically implanted electrode array does. But the crucial reality of the BCI field today is that *no one actually knows where that non-invasive ceiling is*, and we have barely begun to climb the ladder.

These ladders illustrate the technological milestones BCI can achieve as it matures:

Emotional and physiological biomarkers. Your device could record and monitor your emotional state continuously. Knowing when you are happy, sad, angry, stressed, etc.... We are already

scratching the surface of this today.

Attention and focus tracking. The device knows when you are in deep thoughts versus when your mind has wandered. Tracking focus could help improve the ability to do deep work and has numerous health applications. This is also starting to become possible.

Sub-vocal text. This is the rung that changes daily life. Almost-mouth a word, and watch it appear on the screen before your jaw has finished moving with a lag below the threshold of human notice. This allows people to work hours without touching the keyboard once. We already have early lab instances of this technology.

Once you reach higher rungs, things get stranger.

Silent inner-speech recognition. The words you are thinking, decoded without any muscle activation at all. You compose a message in a meeting and nobody around you sees a single muscle move. There are very early, successful lab demonstrations of inner-speech decoding.

Visual concept reconstruction. Hold an image in your mind's eye and your phone shows you a still of it. You can sketch by remembering. Surprisingly, there are some early successes here too; this is less sci-fi than it might seem.

Multi-modal imagination decoding. Not just a still image anymore, the full imagined scene with sound and motion. What a memory feels like to you could be exported with reasonable fidelity to a file that other people can open.

Episodic memory replay. A recording, in your own native cognitive format, of an hour from last Saturday, which could be played back by you as cleanly as you played it back from inside your own head.

Continuous semantic transcription. A real-time transcript of your thoughts. Or in other words: mind reading. The moment this becomes technically possible, every privacy question surrounding this topic will come forward at once. Before this rung can ever reach a consumer, the field has to design an entirely new kind of social firewall, because navigating the societal implications of mind reading will be vastly harder than writing the algorithm itself at this stage.

The top of the ladder is something we do not yet have a clean word for: **a full real-time bidirectional channel** between human cognition and machine reasoning, where the boundary between “what I thought” and “what we computed together” stops being legible. This is what some call the Merge.

The point of the ladder is to illustrate how far a non-invasive BCI could get us and what it could unlock. You only need the bottom rung to ship at consumer quality to have the rest follow through the [flywheel effect](#) ▶. The first wave of non-invasive BCI will likely look like something modestly useful, probably wellness wearables, and then from there, every generation will have exponential gains in capabilities. We are now standing at the edge of the first wave.

The Calibration Ladder

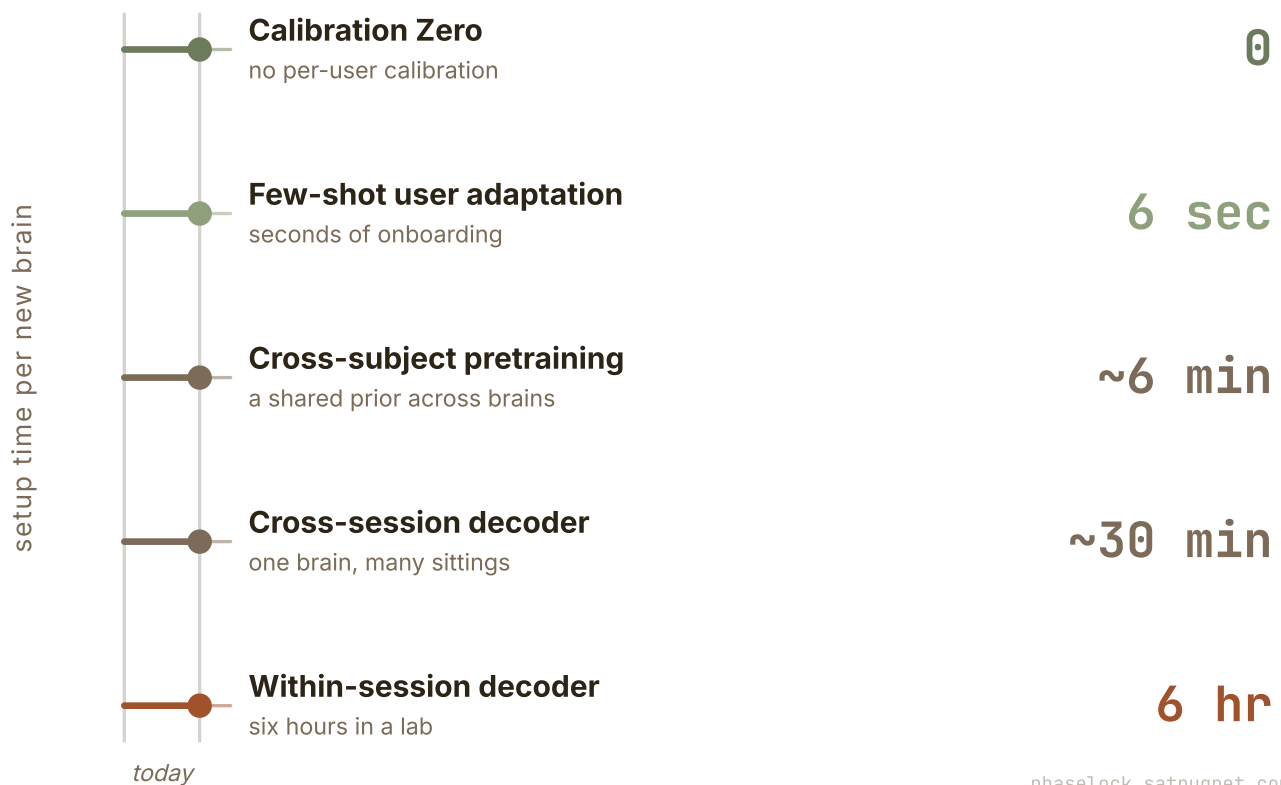
But there is a hidden catch. The capability ladder assumes the device already understands *you*.

In reality, brains are like fingerprints, no two are wired exactly the same. That means every single brain has to be decoded differently. Right now, a BCI is like a custom-tailored suit; it only fits the exact person it was measured for. Before a device can decode your thoughts, there is a calibration period where the model has to learn your specific neural “accent”.

This added friction is the single biggest bottleneck to building a consumer device. In a lab, a graduate student might be willing to patiently train a decoder for six hours but, in the real world, a consumer has zero patience. They will throw the device in a drawer if it does not work almost immediately.

That brings us to the second ladder, running along the capability one, and arguably more important for having non-invasive BCI used outside of the lab. I call it the **calibration ladder**; how much effort each new user has to invest before the device works for them.

THE CALIBRATION LADDER



To create a product, decoding one brain once is not enough; we need to be able to decode a completely new brain cheaply. On top of the calibration, the model has to run fast enough that the user does not notice the decoding delay; optimally with a latency below the threshold of conscious wait.

The capability ladder is *what* the device can read; the calibration ladder is *how much friction* the read has.

The expectation is that pretraining across many brains collapses the marginal cost of decoding a new brain, exactly the way zero-shot learning solved this exact problem for language models.

The strategy described in this essay is to scale non-invasive BCI, using the same scaling strategies that AI has used over the past decade, to build a first generation of BCI consumer products in order to **ignite a flywheel effect which will in turn accelerate the BCI field as a whole.**

The biggest unaddressed remaining unknown is whether non-invasive BCI has a high enough ceiling on the capability and calibration ladders to achieve this goal. The honest answer is that the entire field is currently staring into a fog of war.

13 The Unknown Ceiling

“We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run.” — Roy Amara (Amara’s Law)

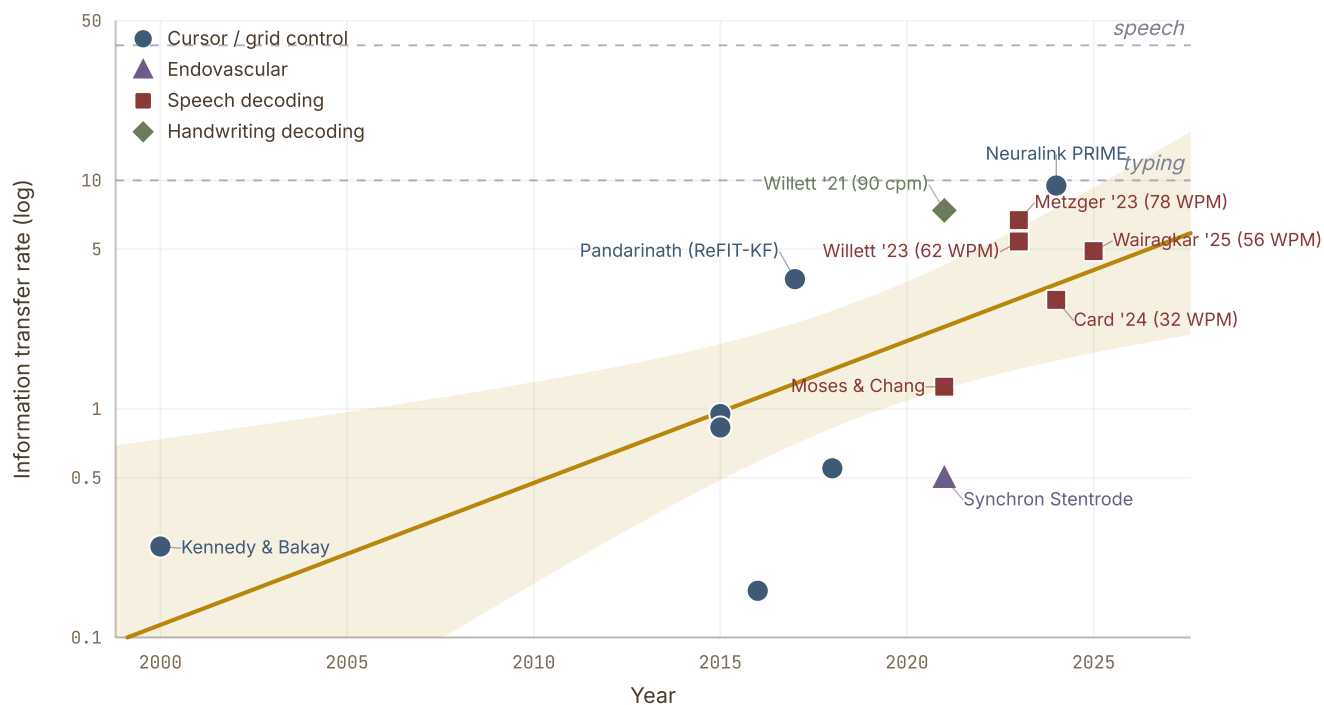
How far does this playbook actually take us? With what timeline? And will the physics even allow it? Well, nobody actually knows.

The historical track record of high-confidence BCI predictions is mostly a graveyard of wrong timelines. I am offering you mine anyway.

The bear case is that the skull, the scalp, and the Earth’s magnetic field combine into a noise floor that may cap how much signal we can pull from outside the head. If that worst case turns out to be true, non-invasive caps out at “modestly useful for medical contexts”, and the only path to high-bandwidth thought decoding is invasive surgery.

But here is the reality on the ground: we are *nowhere near hitting that wall* in actual experiments. The biggest non-invasive datasets are still single-modality, with under 10K hours on just a few dozen subjects, no theoretical argument can override this empirical fact. All the current empirical experiments we have point in the exact same direction: straight up the log-linear scaling law curve, the only way to know where the summit lies is by climbing towards it.

BCI INFORMATION TRANSFER RATE PER YEAR

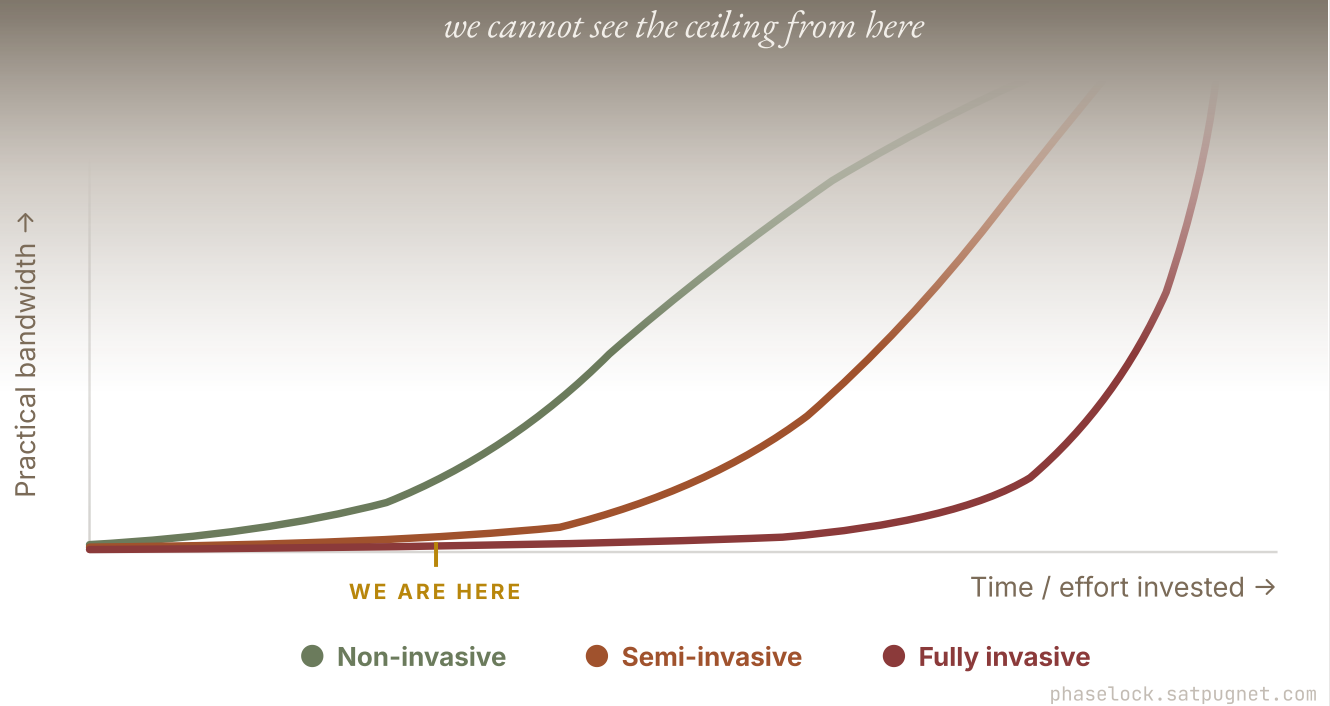


phaseslock.satpugnet.com

Information transfer rate is following an exponential growth, typing and talking speed are now only a few years away. This graph illustrates the trend for invasive BCI, and as seen earlier, similar trends are now visible for semi-invasive and non-invasive.

Earlier in this essay, if you remember [the three S-curves](#) ▶, we illustrated the theoretical upper bounds of what each modality could achieve. This was just an illustration; in practice, we are standing at the absolute **bottom of the curve, looking up into the dark**.

THE WALL IN THE DARK



The [three S-curves](#) ▶ from our current viewpoint. The only way to understand what are the limits of what is possible is to build the ladder and climb it.

For this scaling thesis to work and truly compound, the hardware must clear the friction bar for everyday consumer adoption, and we already have early evidence that this is happening. Just as we saw with language models, **BCI scaling laws are already becoming visible in the lab, and more predictable.**

If we assume the trajectory holds. In the best-case scenario, within three to six years we will see the first ten million non-invasive brain interfaces deployed in the wild. Within six years, they will evolve from early-adopter tech into a ubiquitous layer of everyday computing. The scaling laws will feed on this massive new data stream, creating a *flywheel* that makes the field exponentially more *attractive to capital and talent*.

So where does that actually leave us on our two ladders? How high do we expect to be able to climb before hitting the ceiling?

- **The Capability Ladder:** Inner intent, inner emotion, silent speech recognition, and visual concept reconstruction are all high-probability outcomes without surgery. There are already early signals of this working today. Full multi-modal imagination decoding (exporting a memory with sound and motion) has a low but non-zero probability non-invasively.

- **The Calibration Ladder:** Thanks to foundation models trained on population-scale data, the calibration tax will collapse. Just as modern speech AI instantly understands your voice without calibration the first time it hears you, BCI foundation models will recognize your neural patterns with minimal calibration because they have already seen thousands of other brains. We will reach a point where a new device takes only minutes to adapt to a new brain.

This manifesto relies on the conviction that the ceiling reachable with non-invasive BCI is *much further* away than current consensus assumes. Since theory won't settle it, the only way to verify this conviction is to actually *run the experiment*.

But running the AI playbook requires a massive amount of data, and while for the first few orders of magnitude we can collect the data in the lab, at a certain scale, it becomes intractable. It will only scale if *millions of people* actually wear the hardware. To trigger this rapid scaling loop, the field has to step out of the lab and answer a very practical question: what actually convinces the first ten million everyday consumers to put a brain-computer interface on their heads?

14 The Consumer Wedge

“Any sufficiently advanced technology is indistinguishable from magic.” — Arthur C. Clarke

We have spent a lot of time on theory, but it helps to ground our thoughts in reality. And even if we alluded to it, we have not yet clearly illustrated what the first wave of non-invasive BCI products could actually look like.

In [Section 9](#) ▶, we stared down a four-order-of-magnitude gap between the data AI uses for frontier models and the data BCI currently has. To close this gap, we need orders of magnitude more users, millions of people using BCI on a regular basis.

This raises an obvious question. Who on earth are these **10 million people**, and what exactly are they putting on their heads?

Normal people do not buy abstract scientific milestones, they do not care about “log-linear scaling laws” or “foundation models”, they buy a product because it solves a specific problem or feels like magic.

SIDENOTE

Big tech is already positioning itself for this transition. In 2025 Apple introduced the “Human Interface Device” protocol which allows BCI to register as first-class input on the iPhone.

The wedge that gets 10 million devices onto **10 million heads** is probably *not* going to be one single, *monolithic killer app*. It will be a cluster of “first-good-apps”, products that offer enough immediate,

undeniable utility to overcome the physical friction of strapping a new piece of hardware to your skull. Early adopters might tolerate a clunky wearable just because of the sci-fi novelty, but to hit true consumer scale, the utility still has to be obvious.

Two primary breakout products look most plausible to cross the chasm first.

The first is **frictionless communication**, aka telepathy, talking to machines without using your thumbs or your vocal cords. It would unlock sub-vocal text and intent decoding for smart glasses, coding companions, and AI assistants; making it *faster* than touch, *quieter* than voice, and *ambient* by default. It also provides hands-free augmented reality control (like the Meta Neural Band shape) and forms a critical accessibility bridge for users who cannot speak or type comfortably.

The advantage comes from bandwidth and seamlessness, reading directly from the brain allows us to capture much higher-fidelity intent than a keyboard or microphone ever could. In the beginning, the raw read will not be perfectly accurate, but it does not have to be. Smart product design will mask the early hardware limitations, and flawless, paragraph-length thought-to-text will come later; the first wave will likely rely on high-level intent reading and coarse-grain indicators. As we pair coarse telepathy with an AI that is smart enough to infer the rest of your context, it will start *feeling like magic*²⁵.

The second is **wellness and recovery**, the *Whoop for the brain*. The longevity movement brings intense consumer appetite for neural interfaces. Users have already adopted rings and wristbands to track their physical recovery; a neural wearable provides the missing supervision layer by tracking the central nervous system readiness, emotional states and passive continuous stress tracking. This enables numerous other adjacent applications requiring visibility into the cognitive state such as professional fatigue auditing in high-stakes environments like surgery or air traffic control.

Other plausible first good applications of non-invasive BCI include:

- **Brain-Aware Environments:** Software that is no longer blind to your internal state. An empathetic AI assistant can ingest your cognitive load as part of its prompt context, giving shorter, punchier answers when you are exhausted, providing deeper analogies when it senses you are confused, or entirely suppressing notifications when you are in a flow state. Similarly, study software can track attention and focus, adapting a lesson plan the moment your mind wanders to lunch.
- **The Inner Space Stack:** Tools designed for exploring and optimizing human consciousness. This includes closed-loop meditation apps that track a genuine neural “depth score” and ring a gentle mindfulness bell when focus breaks; sleep and dream products that track REM cycles for lucid-dreaming cues or trigger smart alarms at your lightest sleep moment²⁶; and therapeutic tools for optimizing mental health and clinical depression.

25. Once embedded into consumer devices, BCI will also come with numerous new UX questions: When should the system act on a thought, and how fast? Which thoughts matter, and which are just passing cognition? How to build a thought native UI interface? How to translate different mental mechanisms across people into actions? Overall, this will demand a UX overhaul of our way to interact with the digital world even more significant than the ones the internet and AI brought.

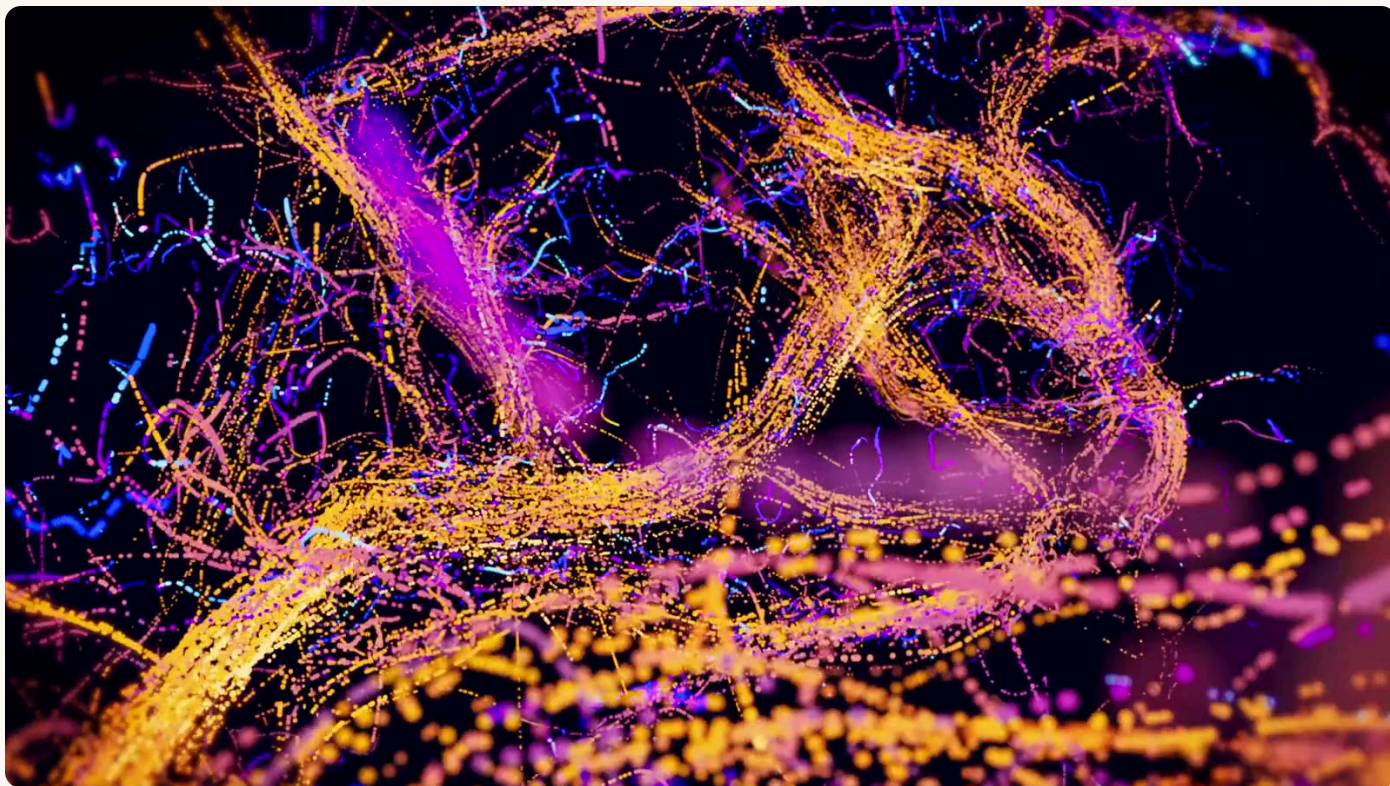
- **The Creative Engine:** Thought-to-image generation as a professional creative tool. Instead of typing complex text prompts into an image generator, a designer holds a visual concept in their mind's eye and watches the system render a rough draft on screen. This translates to music and movie creation as well, any medium that captures a mental state and tries to translate it into creative software could benefit from this.

Once 10 million devices are used in the wild everything accelerates, manufacturing costs collapse, the sensor stack becomes commoditized, data becomes plentiful and the foundation decoder models become shared public goods.

This **consumer flywheel** strategy can shorten the timeline for progress significantly. By shifting the field out of the slow, methodical timeline of academic tenure tracks and clinical trials, non-invasive BCI gains a massive structural advantage, it allows data to accumulate *at scale*. Given the staggering velocity of frontier AI development, this consumer-led explosion is the only path moving fast enough for BCI to have a shot at keeping up with frontier AI, and for humans to remain in the loop.

If BCI, and neurotechnology overall, is going to play a role in keeping humans in the loop during the transition to superintelligence, it has to catch up, and it has to do it fast. It certainly cannot move at the agonizingly slow pace of invasive surgery, which is why the non-invasive consumer loop is the only vector fast enough to outrun an impending AI takeoff within the narrow window left to pull it off.

26. If dream data proves useful at all for training brain decoders, then sleep could become one of the most effective ways to collect neural data at scale.



The dense vascular network of a living brain, traced vessel by vessel. Visualization by [Aleph Neuro](#).

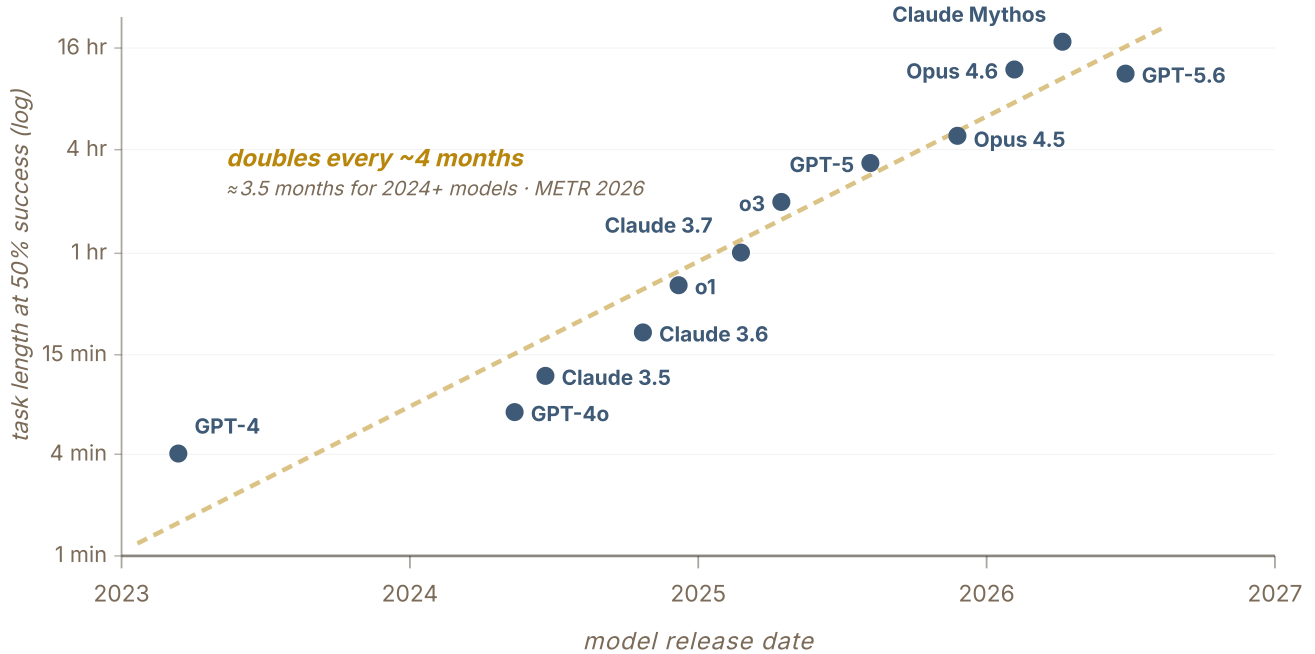
15 Phase Lock

“The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought ...” — J. C. R. Licklider, Man-Computer Symbiosis, 1960

AGI is coming soon. At this point it is no longer a distant thought experiment, we are all expecting the future to arrive early, and that accelerated timeline raises a foundational question: how do we actually keep humans in the loop?

Frontier AI is moving ever *faster*, with flagship models now released monthly, sometimes faster. The window of opportunity to shape what happens next is terrifyingly narrow. **These are the last few years of the world as we know it.** [Metaculus](#), a prediction website with an impressive track record, currently estimates that weak AGI will be reached by mid 2028. That leaves us with a couple of years, at most, before we fully cross the intelligence threshold and enter the *New World*.

AI CAPABILITY DOUBLING TIME



The length of time an AI can operate fully autonomously, on a log scale; it doubles roughly every four months. Source: [METR's benchmark](#).

The critical decisions, what these models learn to optimize, whose preferences carry through, and how they ultimately interface with humanity, are being locked in right now. If our goal is to maintain agency and actually keep humans in the loop as intelligence keeps improving, human-machine interfaces need to arrive *within the next few years*.

I call this brief interval between humanity recognizing that AI will surpass us and the moment it does the **Phase Lock Window**: the narrow span in which we still have the agency to shape the relationship between human and machine intelligence before it hardens into something far more difficult to change. The Phase Lock Window is roughly between now and the mid-2030s.

SIDENOTE

To be clear, BCI will not magically solve every problem raised by AGI. There is a mountain of socio-economic, philosophical, and technical alignment problems we also have to figure out; and the technology, as [Section 22](#) ▶ lays out, is itself very dangerous. However, **BCI, and neurotechnology, are the only structural fix to the human-machine communication and cognition bottleneck**, the sole solution expanding human capabilities rather than constraining machines. This makes neurotechnology one of the strongest long term levers for AI alignment. But it can only play that role if it scales in time, we only have a few years to make BCI part of the solution.

That urgency is why this manifesto proposes a concrete strategy for accelerating the field through non-invasive BCI. If establishing a high bandwidth brain-computer interface requires invasive surgery, it will *not* land this decade. And by the time it does, the transition to artificial superintelligence will have passed, and most of the questions worth asking will have already been answered using whatever limited tools we had on hand. Because the surgical route cannot close the data gap inside the Phase Lock Window, the explosive adoption of non-invasive hardware is our only remaining vector to integrate biology into the alignment toolbox before the future is permanently locked in.

When mapped against the arrival of superintelligence, this consumer-led BCI explosion is not just a cool product milestone, it is more akin to an urgent rescue mission

Blueprint and Convergence

Throughout [Part II](#) ▶, we have asked one core question: *Does the AI scaling playbook work for the brain?*

The answer, as far as the evidence can take us for non-invasive BCI, is yes.

When you strip away the messy biology and treat the brain as a computer science problem, the similarities to the early days of the AI era are *flagrant*. The data speaks for itself; BCI has the same footprint as the early deep learning boom, with an eight-year lag (see [the eight-year-offset chart](#) ▶).

We now have a complete, cohesive blueprint for scaling non-invasive BCI. The final takeaway of [Part II](#) ▶ is that we are standing in a short, weird and fragile window where BCI and neurotechnology can still matter for shaping the future of humanity, but this Phase Lock Window is remarkably narrow.

Yet scaling AI and BCI in isolation tells only half the story. To understand why BCI and neurotechnology as a whole become more important as AI grows more capable, we must examine how the two technologies converge. Neurotechnology inspired modern AI; AI now accelerates neurotechnology; and as BCI matures, it will become foundational to AI once again. BCI is about to become the infrastructure we use to both *build* and *operate* frontier models, providing the multimodal neural data needed to inform the training and alignment of the next-generation of AI models, and creating a high-bandwidth interface to control them directly through thought rather than words. Crucially, this bridge is one of the few alignment levers capable of scaling alongside AI itself.

AI and BCI are now on a direct and imminent collision course.

CHECKPOINT II

Zooming out, this is where we stand.

In [Part I](#) ▶, we asked ourselves: *Can the brain actually be read?*

In [Part II](#) ▶ we took the AI playbook and asked: *what happens when you run it on the brain?* We explored how applying AI's scaling strategy to non-invasive BCI could unlock the first generation of BCI consumer products.

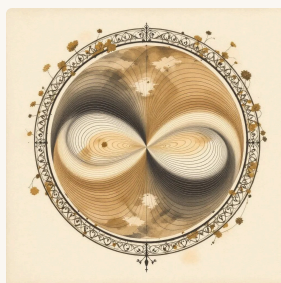
To answer that, we built a four-layer pyramid, *hardware* at the base, then *data*, then *foundation models*, then *product* at the top, and climbed it from the ground up. The hardware layer has to shift from lab-fit to scale-fit. The data layer has to close a four-order-of-magnitude gap with frontier AI corpora. The foundation-model layer has to inherit the open-collaboration and scaling culture that made language models take off.

Then we stepped back from the layers and looked at the system as a whole. We discussed the coordination problem that has kept the field stuck, the cold-start trap where every layer is rationally waiting for the one below to move first. We mapped the two ladders through which to measure progress: the *capability ladder* (what a device can read) and the *calibration ladder* (how cheap it is to make it work for a new brain). We addressed the bear case, that the skull might cap the signal earlier than we think, and concluded that no whiteboard argument can settle the question; the only way to find the ceiling of non-invasive BCI is to empirically test its limits. And finally, we looked into the two consumer wedges, *frictionless communication* and *continuous wellness tracking*, that have the best chance of sneaking this hardware out of the lab and onto ten million heads.

Along the way, I have made three fundamental assumptions:

1. **Non-invasive brain signals carry enough information.** The brain needs to emit enough information for a sensor outside of the skull to be able to decode it. Early results suggest they do.
2. **Scaling laws hold for brain data.** More data must continue improving performance. The exponential curve reported by [Banville et al.](#) ▶ must persist at 100K-hour and 1M-hour corpora rather than plateau early.
3. **Brain foundation models generalize across people.** Models trained across enough subjects must reduce per-user calibration from hours to minutes or eliminate it entirely.

If all three hold, **telepathic brain-machine communication will become a consumer reality this decade**, setting in motion a powerful flywheel that accelerates the field and helps neurotechnology close the gap with AI.



PART III.

The Collision

Why does BCI matter for AI?

16 Converging Trajectories

“Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended.” — Vernor Vinge, 1993

What does it actually mean to be on a collision course, and why does it matter so much? After all, AI and BCI have both been around for more than 50 years. Shouldn't they already be moving along predictable, well-worn tracks?

Maybe in the *Old World*, but that world ended when frontier LLMs arrived and AGI became a possibility. Like everything else in the technological biosphere, BCI is no longer evolving on its old trajectory, **the gravitational pull of AGI has altered its course.**

Historically, prior to AI, technologies lived in fairly compartmentalized worlds. Biotech had little to do with spacetechnology, Taiwan had little to do with SaaS valuations, and robotics had little to do with accounting; every field had its own box.

But AI has melted the walls between the boxes.

The reason is simple, every field is secretly made of the same things: decisions, predictions, search, design, optimization, communication and labor. AI can operate on all of that, so it does not behave like a new industry, it behaves like a new amorphous layer underneath every industry.

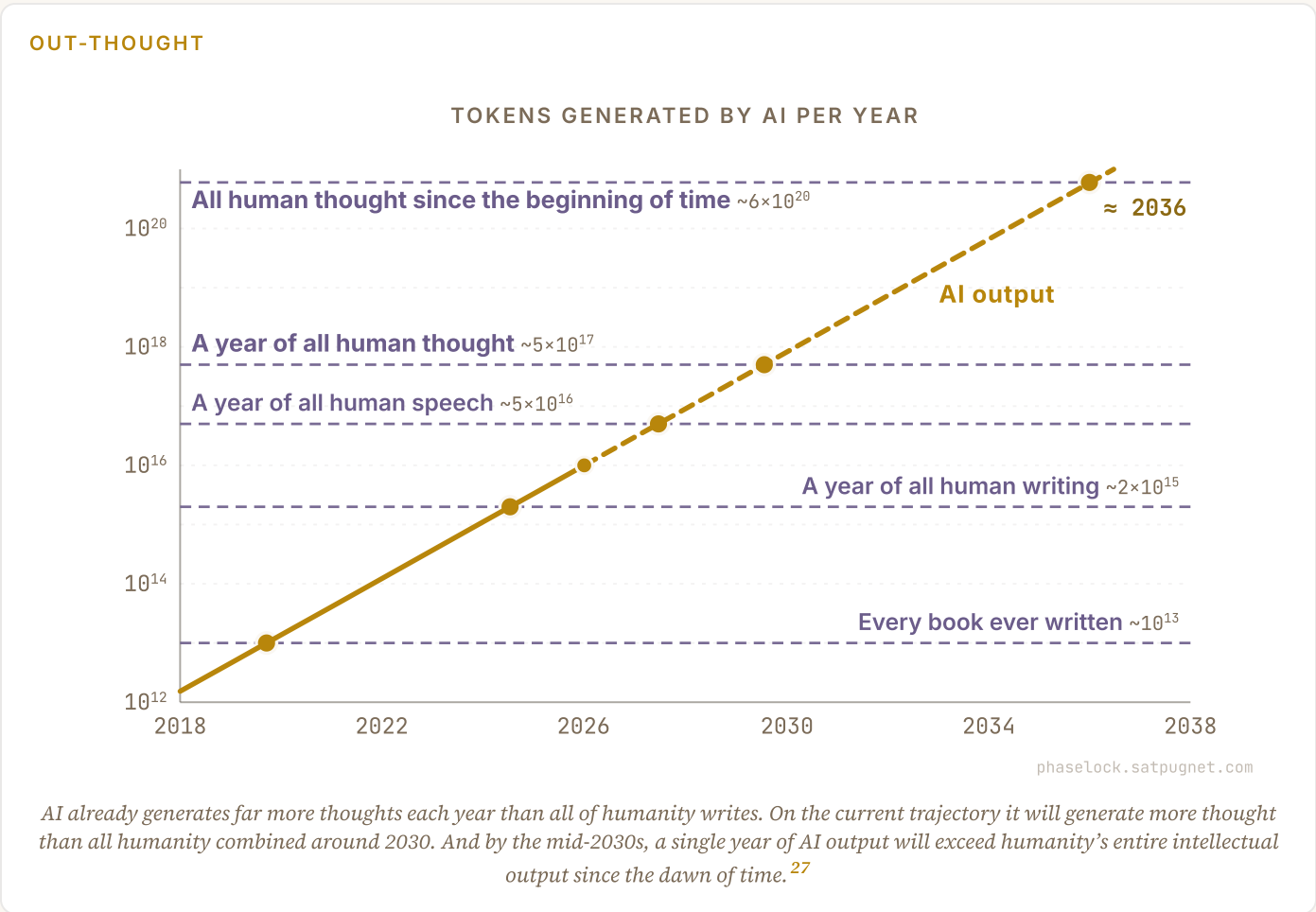
That is why AI is not just “*another technology*”. Chips made computers faster, but chips did not sit around trying to invent the next generation of chips. Electricity powered factories, but electricity did not redesign the factory. The internet connected people, but the internet did not decide what should be built next. Humans did the thinking. The tools amplified us, but they did not replace the engine of iteration.

The deeper discontinuity is that artificial intelligence is no longer just accelerating the work; it is actively taking over the engine of iteration, especially when paired with robotics. It can write code, design experiments, interpret results, propose next steps, and eventually act in the physical world; and it will not politely stop every few weeks to ask whether humanity is emotionally ready for the next order of magnitude. It will just keep optimizing and happily self improve into oblivion.

That is the scary and important part: we are crossing from passive technology to participatory technology. From tools that wait for human minds to push the world forward, to tools that can increasingly help push the world forward on their own. It is the first time in history that a self improving technology is created.

On the current trajectory, humans will be pushed out of the loop within one decade, maximum two, and probably sooner.

I do not necessarily mean humans disappear. I mean something more subtle and insidious: the progress that drives civilization will begin running at machine speed. Research, engineering, strategy, governance, capital allocation, warfare, logistics, and eventually physical production will increasingly become AI-mediated, AI-accelerated, and AI-executed. Humans will remain present, but no longer central.



That may be acceptable if the direction AI is driving the world toward is fundamentally aligned with human flourishing, and with the flourishing of the rest of our conscious neighbors. It gets less fun if the system's emerging dynamics **lock civilization into something brittle, alien, or dystopian** before we can meaningfully intervene.

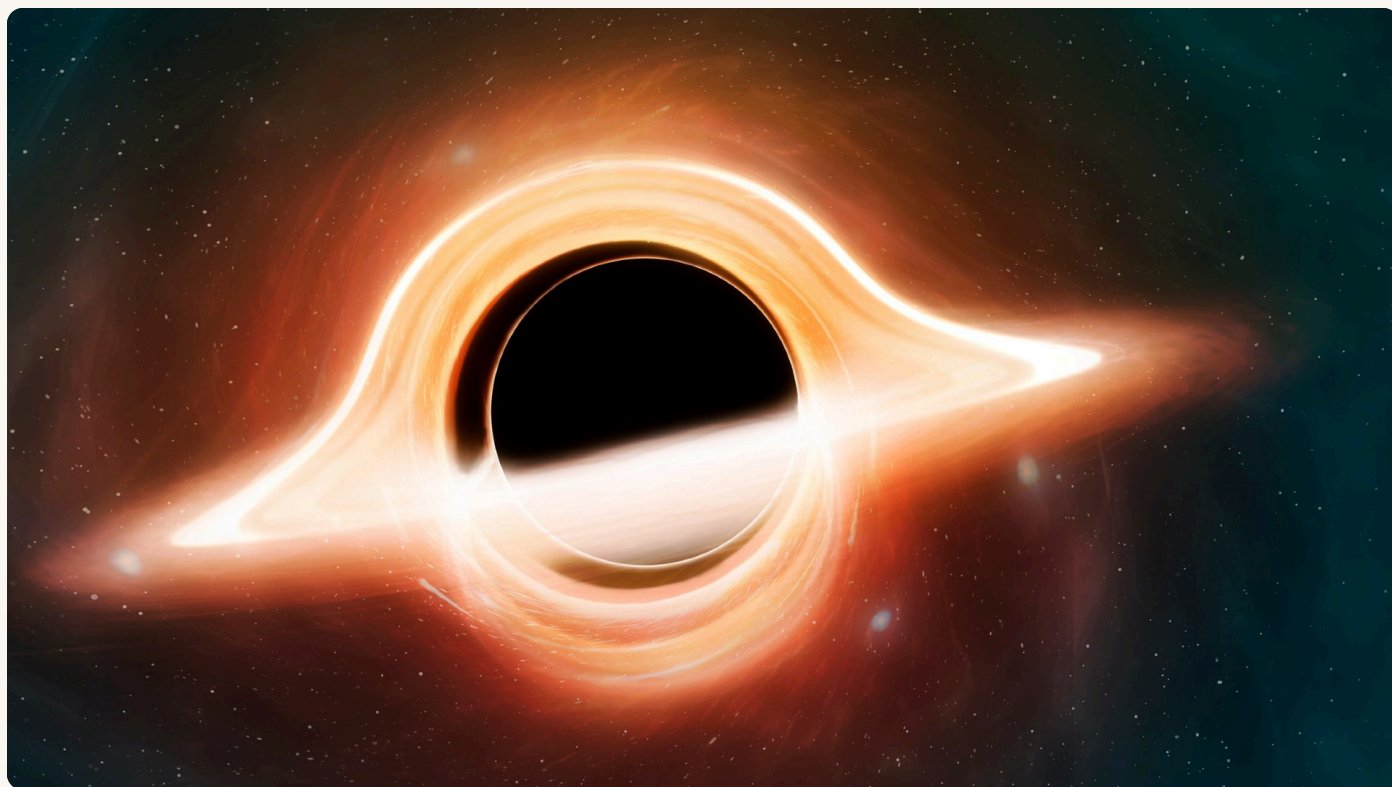
Nobody knows exactly where this goes. But right now, the variance of possible futures is higher than it has been at any point in human history. Thirty years from now, we could plausibly be living in worlds ranging from almost unimaginable abundance to forms of dystopia that previous generations did not even have the conceptual vocabulary to fear.

27. Estimates based on [OpenAI](#), [Google](#), and [words-per-day studies](#).

If the timeline to AGI feels too sci-fi to internalize, consider how you viewed the future in 2020. You likely imagined slightly faster computers, sleeker electric cars, or perhaps cleaner energy grids. I would bet you one BCI headset that you did not envision a world where a machine could draft your emails, write your code, tutor your children, summarize scientific papers, generate videos, handle large parts of your daily work or solve century-old math problems.²⁸

Now apply the same correction forward. Imagine you are making the *same* mistake today, mispredicting the rate of improvement in exactly the same way. If you are using the same linear intuition that failed to predict today to predict the next ten years, you should be suspicious of the answer it gives you. This little mental experiment is a useful way to fight the brain's instinctive blindness to exponentials.

The goal of this experiment is not to shame you for having a normal human brain. In our defense, humans have rarely, if ever, had to think clearly about changes this fast. Our brains were not built to intuit exponentials.



Just like a black hole, AI distorts reality around it.

But why does that matter for BCI? Or more importantly, why does BCI matter for AI?

²⁸. Literally as I write those lines, OpenAI just released [10 Math breakthroughs](#) done by ChatGPT alone.

Because it is one of the few plausible mechanisms for humanity to retain agency, one of the few mechanisms to push the only long term solution we have: *alignment-by-integration*.

As we enter the *Phase Lock Window*, the collision between AI and BCI will play out across three main surfaces:

The first surface is training. Neural data will become the next major modality for AI and robotic labs. It is denser than text, video, or any modality scaled to date, and structurally upstream of all of them: the raw, unedited signal of human cognition rather than just an artifact of it. [Section 17, *The New Major Training Modality*](#) ▶, examines why robotic and frontier labs would actually buy a neural corpus and unpacks the math behind neural data as a training modality.

The second surface is interfacing. High-bandwidth BCIs will become the default mechanism for AI-human interaction; keyboard, mouse, and feedback channels rolled into a single continuous stream. [Section 18, *The New Interface*](#) ▶, shows how sub-vocal input, real-time intent decoding, and continuous cognitive context will create a new way to interface with AI, and why communication bandwidth has increasingly become the binding constraint for AI usability.

The third surface is alignment. Neural signals will become one of the few alignment levers capable of scaling alongside frontier AI capability, closing the bandwidth gap between humans and machines to allow humans to remain meaningfully coupled with the system. [Section 19, *Why This Matters for AI Alignment*](#) ▶, makes the case for BCI, and neurotechnology, as one of civilization's main shots at alignment as AI capability blows past human supervision speed. It does so in three ways: by providing a high-fidelity reward signal for RLHF, by providing a much better understanding of how human values are encoded in the human brain, and ultimately through human-AI symbiosis and by upgrading our biological minds.

AI and BCI are now converging from opposite directions. Artificial neural networks began as abstractions of biological neurons, but AI quickly scaled far beyond neuroscience. Now their relationship is becoming reciprocal: AI is accelerating our ability to decode the brain, while BCI is emerging as a way to train, interface with, and align increasingly capable AI systems.

The rest of [Part III](#) ▶ walks through each of these three surfaces in turn, starting with the new training modality BCI unlocks for AI.

17 The New Major Training Modality

“we can know more than we can tell.” — Michael Polanyi, *The Tacit Dimension*, 1966

If all the text on **the internet were one mile long**, it would only take **one year of thought for 50,000 people to reach the Moon.**²⁹

And yet, barely anything from this *sea of thoughts* ever leaves our minds.

But this won’t hold true for long, brain data is about to become a new major training modality for frontier models.

This certainly sounds crazy at first, but let me unwrap how I reached this conclusion through a first-principles analysis.

To understand why a frontier AI lab would care about the electrical noise inside your skull, you have to look at what they have been ingesting for the last five years. And we have to look past the current paradigm of data collection to ask ourselves: *what is AI training data, really?*

The Map Is Not the Territory (and Text Is Not Thought)

The history of modern AI is essentially a story of data scavenging. We took the artifacts humans left behind on blogs, books, transcripts, the internet, and videos that make up the messy, digitized residue of civilization, and used them to reverse-engineer the underlying logic of human intelligence. The story of modern AI, since the Transformer, has been driven by data, not algorithms; the models’ architecture barely changed along the way.

Indeed, AIs has been ingesting an expanding sensory diet, each step unlocking a richer slice of reality for the models. **Text** gave them compressed human knowledge. **Code** gave them executable reasoning and a training ground for strict logic. **Images** gave them perception. **Audio** gave them speech and timing. **And video** gave them motion and causality unfolding through time.

TEXT ► CODE ► IMAGE ► AUDIO ► VIDEO ► **NEURAL**
phaseLock.satpugnet.com

But there is a massive, systemic gap in this approach. Every dataset we currently use is a highly compressed, lossy, downstream byproduct of the actual cognitive engine: our brain.

²⁹. Fermi estimate: 16 billion cortical neurons × 0.16 spikes per second × one assumed bit per spike, for 16 hours daily; compared with approximately 1.2 petabytes of usable public internet text.

When you write a sentence, record a video, or write a line of code, your brain has already performed an immense amount of high-dimensional computing. What hits the screen or keyboard is just the tiny, serialized exhaust of that process.

- **Text is not thought.** It is a low-bandwidth representation of a conclusion, stripped of the hesitation, the counterfactual paths explored, and the mental architecture that built it.
- **Code is not cognition.** It is the final execution script, missing the entire invisible debugging process and structural rationale that occurred in the engineer's mind.
- **Speech is not intent.** It is the acoustic compression of an internal state, constrained by vocabulary and vocal mechanics.
- **Images or videos are not experience.** They capture the external photons bouncing off an object, not the internal world model or spatial intuition of the observer.

Neural data is not just another dataset to add to the pile, **neural data is structurally upstream of all other modalities.** It is the raw, unedited, high-fidelity signal of human cognition itself. It sits upstream of basic mechanics like perception, attention, and memory; upstream of active outputs like language and motor action; and crucially, upstream of the complex, invisible machinery of moral judgment, aesthetic preference, and creative agency. By training models directly on brain data, AI labs would stop studying the shadows on the cave wall and start looking directly at the fire.³⁰

³⁰. Plato's Allegory of the Cave: prisoners chained facing a blank wall see only the shadows of objects passing behind them, and mistake the shadows for reality. Here the artifacts we train on (text, video, code) are the shadows; the neural activity that casts them is the fire.



One of the first times a thought moved a cursor across a screen. BrainGate, 2004.

Exposing the Tacit Knowledge

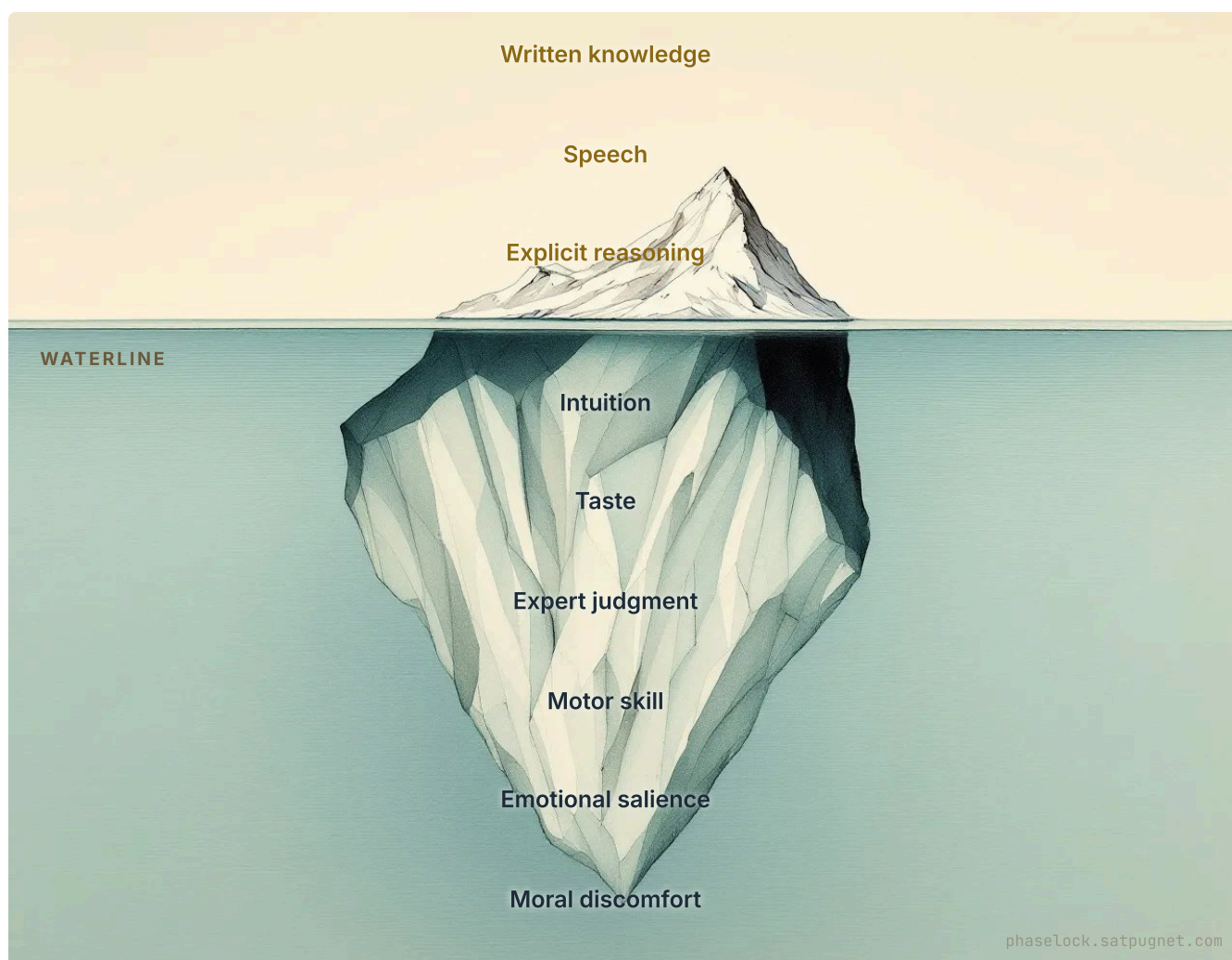
This structural gap cannot be fixed by just asking humans to document their reasoning. A staggering amount of human knowledge simply cannot be captured by a keyboard or a camera because it

operates entirely outside of conscious language. Economists and philosophers call this *tacit knowledge*: the things we know how to do but cannot tell how to do.

Think about the following cognitive states:

- An expert mathematician's sudden **aesthetic intuition or scientific hunch** that a specific proof vector “feels right”.
- A seasoned surgeon's instantaneous **pattern recognition** during a high-stakes crisis.
- The profound **moral discomfort** or subtle cognitive friction experienced when an AI output contains a deeply hidden ethical anomaly.
- The raw **mechanical feedback loop** of an elite athlete or pilot navigating a complex trajectory.

THE TACIT-KNOWLEDGE ICEBERG



Text captures the visible tip, but the mass that actually drives expertise never reaches a keyboard.

You cannot scrape this knowledge from the web because it does not exist on the web, it cannot be articulated in words. If you ask a top researcher how they came up with a breakthrough hypothesis,

they will often point to a vague “hunch” or “intuition”.

That hunch is not magic; it is a physical computation happening inside a biological neural network. For the first time in history, non-invasive consumer BCI gives us a net to catch these fleeting, unrecordable signals. Collecting an expert neural corpus allows frontier labs to train models on the purest form of human expertise, capturing the unspoken, unwritten dark matter of human intelligence.

Those cognitive processes can be inferred to some extent of course, and this is what current models have been doing. But reconstructing human reasoning from the artifact of civilization is akin to teaching a blind man to see using nothing but words. It is nothing short of a miracle that it works as well as it does. Nevertheless, inferring is not the same as observing, and it has its limitations; brain data allows the paradigm to shift from inference to direct observation, finally providing direct access to the biological reasoning engine AIs are trying to replicate.

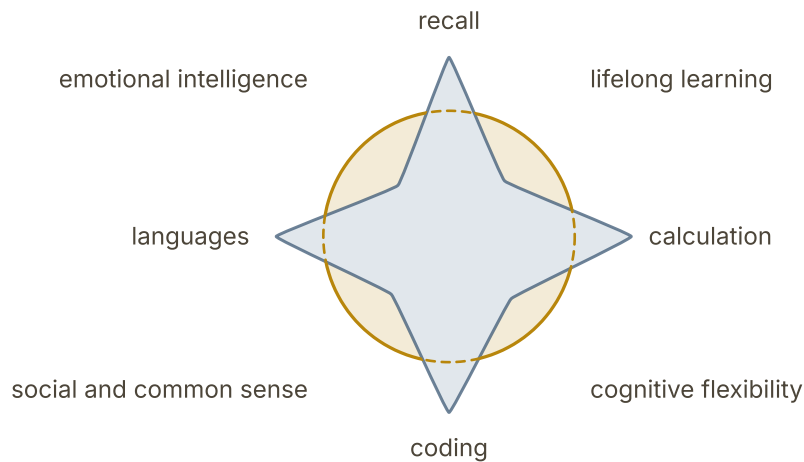
The jagged intelligence of modern AI systems arises exactly from this missing training signal, the **cognitive dark matter** of the mind, brain functions that meaningfully shape behavior yet are hard to infer from behavior alone. BCI datasets could help instill these capabilities into AI systems and produce models with a more human and less jagged intelligence.³¹

31. This [paper](#) ▶ does a great job at discussing this idea in depth.

THE JAGGED INTELLIGENCE FRONTIER

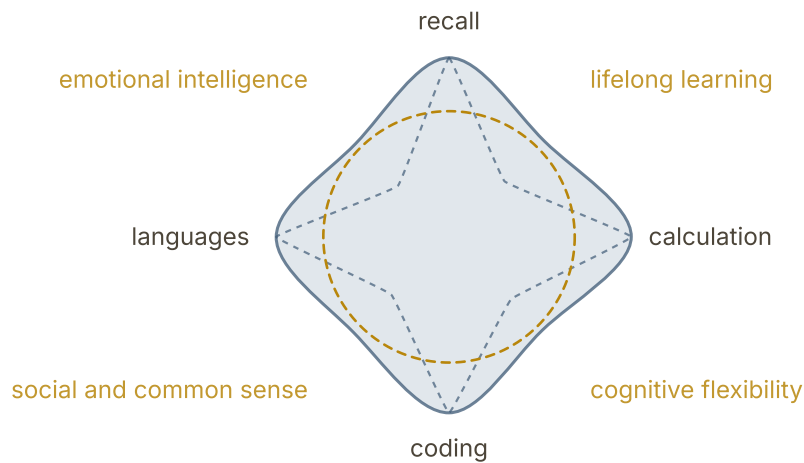
TRAINED ON BEHAVIOURAL DATA

superhuman where benchmarks can score it, far below humans on the cognitive dark matter



+ NEURAL DATA

the dips rise toward human level and the frontier smooths out



✧ AI capabilities ○ human capabilities -- AI without neural data

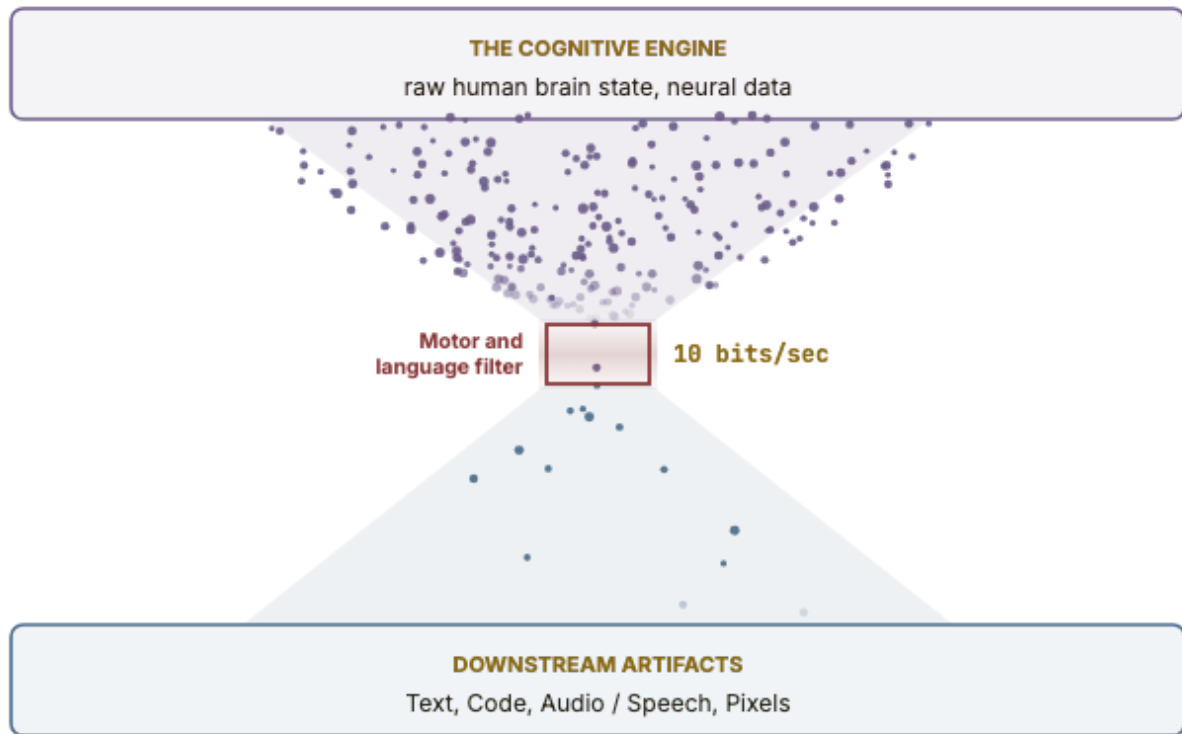
phaselock.satpugnet.com

Neural data can help smoothen the intelligence frontier of AI models by providing direct access to tacit knowledge which is not present in current training data.

Structurally Upstream of All Datasets

Neural data is not just another dataset to add to the pile. It is the raw substrate from which all other datasets are manufactured: the uncompressed, unedited signal of human cognition before it hits a motor or language filter.

THE COGNITIVE ENGINE



Every dataset we train on today is the narrow exhaust of the much wider state above it.

Because it sits at the absolute fountainhead of intelligence, a robust neural corpus allows us to directly observe cognitive properties that are harder to capture through current training modalities:

- **Attention and focus.** The exact millisecond-by-millisecond shift in what a human mind deems salient.
- **Error detection.** The immediate, involuntary recognition that something is broken or incorrect before a manual fix can even be planned. Already used to correct robot mistakes in real time ([Salazar-Gomez et al., ICRA 2017](#) ▶).
- **Preference and judgment.** The subterranean shifts of aesthetic, moral, and social alignment that dictate human choice long before an explanation is post-rationalized. Just by reading brain signals, we are already able to tell with 70–80% certainty whether someone would buy a car ([Tusche, Bode & Haynes, J. Neurosci. 2010](#) ▶).
- **Language and syntax.** The pre-verbal conceptual maps that exist before words are chosen. Researchers have already successfully decoded silent, unuttered inner speech directly from the cortex ([Willett et al., Cell, August 2025](#) ▶).
- **Reasoning.** The hidden search process before the answer appears: confusion, hypothesis testing, cognitive load, uncertainty, conflict, and the moment a solution snaps into place.
- **Intent.** The raw goal a user is trying to accomplish before it is compressed into a physical action.

- **Semantic matching.** The underlying, multi-dimensional geometry of how a biological mind actually carves up reality. We already have proof that fine-tuning language models to predict human brain activity measurably improves their downstream performance ([Schwartz, Toneva & Wehbe, NeurIPS 2019](#) ▶).

By integrating brain data into frontier models, labs can stop trying to reconstruct the human mind by analyzing its footprints, and they can begin training models directly on the machinery that produces cognition: with attention before speech, intention before action, preference before explanation, error detection before correction, and intuition before argument.

Neural data will become the new major modality. Despite brain data being noisy, private, hard to collect, hard to label, and full of traps, it is invaluable at scale, and it is the source of all other types of data we currently use for training. Before a person writes a sentence, clicks a button, rates a response, changes their mind, or notices an error, there is a cognitive state underneath. Today's models mostly see the residue of that state, neural data will be the first training signal that samples the state itself. So far, it is like we trained the models on subtitles versus the audio they came from. Subtitles preserve the words, but the audio preserves the hesitations, the emphasis, and the half-formed thoughts that got compressed out before any line of text was committed.

*Today's models learn from the subtitles. Neural data is
the audio they came from*

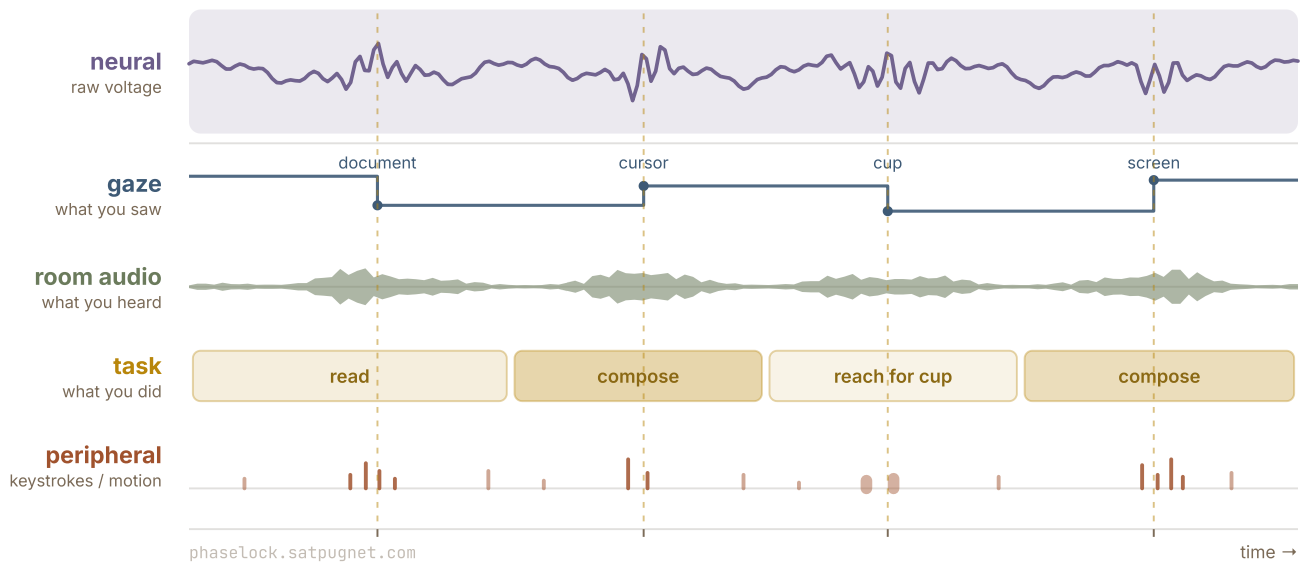
A New Mechanism

We have been talking abstractly about neural data for a little while now without going into the heart of it. It might be a good time to get more concrete. If you are an AI researcher sitting at a GPU cluster at OpenAI or Anthropic, *what does a neural dataset actually look like?* Are we literally feeding squiggly EEG lines into a multi-modal transformer?

The answer is yes, with a few caveats. Raw brain data is not particularly useful on its own. A raw neural recording, a fluctuating voltage or a smear of blood oxygenation in the brain, is close to useless. Hand the smartest model on Earth a terabyte of raw brainwaves with no context, and it will learn only very little. It is mainly *biological noise*.

Neural data only becomes a multi-billion-dollar training asset when it is **perfectly synchronized with environmental context**. A sudden voltage spike in the brain means nothing. But a voltage spike that is time-stamped to exactly what the person was looking at, the sounds in the room, the task they were trying to achieve, and the sensory input hitting their peripheral nervous system, that becomes a wildly valuable training example. The environment acts as the *anchor* that allows the model to understand what the brain was actually doing.

ANCHORED NEURAL DATA



A raw neural trace is close to noise, it needs to be paired with the environmental context to be valuable.

In the short term, BCI labs will likely help AI models digest this information by extracting processed features from the raw physics. Specialized models will translate raw EEG or fNIRS into latent vectors representing attention, error recognition, surprise, or cognitive load. But the history of deep learning tells us that **raw, uncompressed data usually wins out over human-engineered features**. Ultimately, the most valuable dataset is the rawest possible biological signal (EEG, MEG, and the like) paired directly with the stimulus, the task context, and the behavioral output, tracked longitudinally from the same people over months and years.

Ingesting brain data will also certainly require an architecture update to the models. For a modern LLM to ingest this continuous, multi-channel biological signal, it needs a **new ingestion layer**, similar to the architectural lift that was required to go from text-only models to vision-language models. This requires a meaningful amount of research work, but labs do not have to build it from scratch, they can drop a BCI company's decoder in as their neural front-end, before building better architecture on their own. Additionally, because neural signals vary far more across people than text, this ingestion layer may remain *partly personalized*. Each user could have a **brain-to-embedding adapter** fine tuned on their specific neural "accent" and able to plug into any compatible foundation model, this would create a new infrastructure layer and industry of portable, composable neural decoders between brains and AI. Overall, in the context of a multi-trillion-dollar race for AGI, building this translation layer is just a *minor plumbing detail* relative to what it unlocks.

Another particularity of non-invasive brain data is that it is very noisy. Only scale can make the whole thing tractable as **scale launders noise**. A neural signal that is only 55% reliable in one person for one

second is close to worthless, but it is an enormous statistical edge when you have a billion such seconds.

This also brings us back to why **non-invasive portable wearables** are really the only path forward in the foreseeable future. A surgical implant or an fMRI scanner gives you pristine signal, but only from a few dozen patients. The exact data shape the labs need, tens of thousands of hours of naturalistic, context-rich recording drawn from hundreds of thousands of ordinary lives, is one only portable non-invasive BCI technology can produce.

You will likely see brain data integrate into frontier models in a smooth evolution. It will start with **brain-derived labels**, and eventually, as the decoding layers mature, neural data will just become a **native modality in the pre-training mix** itself. This will allow the model to learn the deep statistical relationships between the internal cognitive states that precede human action and the physical world. Eventually, labs will start producing their own *thought-to-text*, *thought-to-speech* and *thought-to-robotic-actuation* models.

SIDENOTE

There is one thing we must explicitly mark before we continue: the very architecture that makes neural data the ultimate training modality is also what makes it the *ultimate instrument of surveillance*. A dataset this valuable is, by definition, the most intimate record of human cognition ever assembled. Just like many technologies out there, there is a dual-use risk which needs to be duly acknowledged. There is no clever engineering hack or quick mathematical band-aid that completely dissolves this dual-use tension. The ethical blast radius of this technology is significant, and is discussed more in depth in [Section 22](#) ▶.

The Bear Case and the Trillion Dollar Alpha

Let us pause the techno-optimism for a second. If you pitch this thesis to a room full of machine learning researchers today, a good chunk of them will roll their eyes.

The skepticism is understandable: brain data is insanely noisy³², hard to label, and orders of magnitude more expensive to collect than scraping the internet. Why not just rely on synthetic data and behavioral traces?

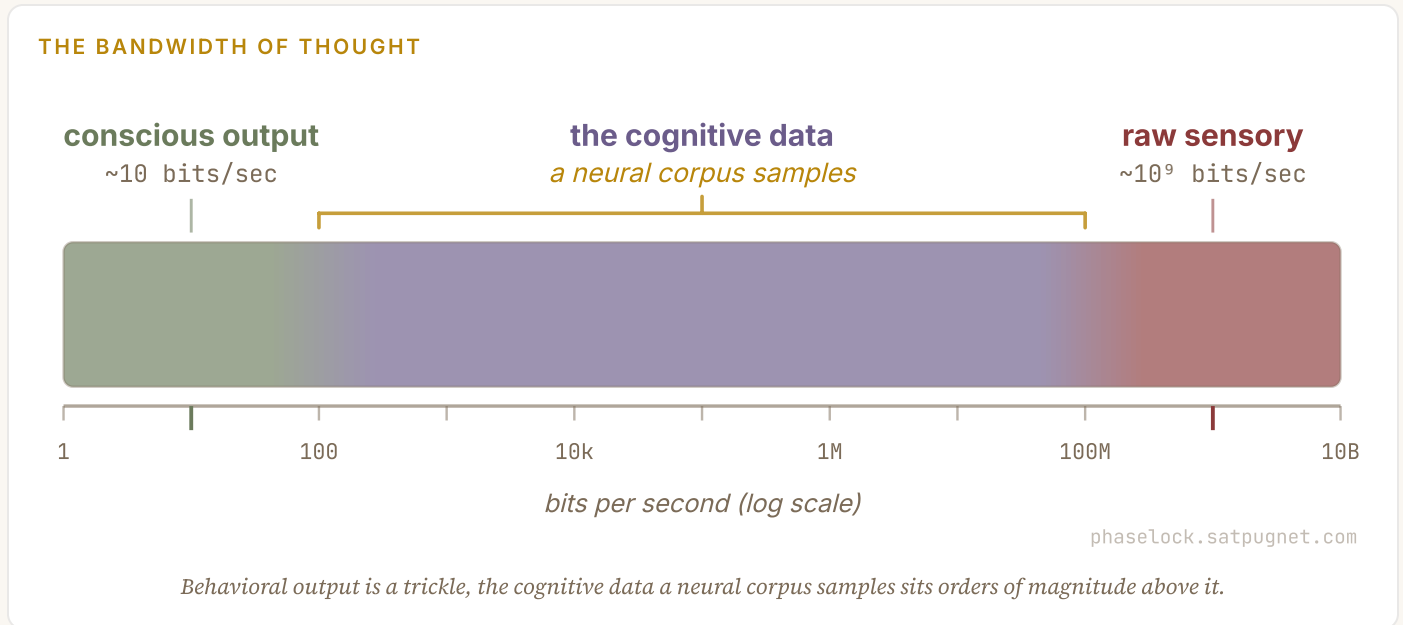
The answer lies both in the **absolute scale of the frontier** and the **potential of this new modality**.

We are rapidly approaching a world where foundation models command multi-trillion dollar valuations. If ingesting a massive, expensive neural corpus improves a frontier model's reasoning or alignment by a mere one percent³³, that single percentage point represents **tens of billions of dollars** in enterprise value. At that scale, labs do not need brain data to be flawless. They just need it to contain a sliver of alpha that their competitors lack.

32. One apt comparison I heard is that decoding non-invasive brain data is like trying to listen to a symphony from outside the concert hall during a hurricane.

33. Early experiments are already showing [two-digit percent improvements in model capabilities](#).

To understand the value of that alpha, consider the sheer volume of information lost between the brain and the keyboard, behavioral output is a *trickle*. The cognitive data a neural corpus samples sits orders of magnitude above it.



SIDENOTE

Now, to be clear, there are not a billion bits of independent, useful “thought” pumping through your head every second. A massive chunk of that 10⁹ is biological noise, overlapping sensory inputs, and redundant physiological signaling, it would be absurd to treat every bit as a discrete piece of training data. But the true, usable cognitive state, the intent, the hesitation, the intuition, sits somewhere in that massive gap between the high-dimensional internal computation and the low-dimensional behavioral output.

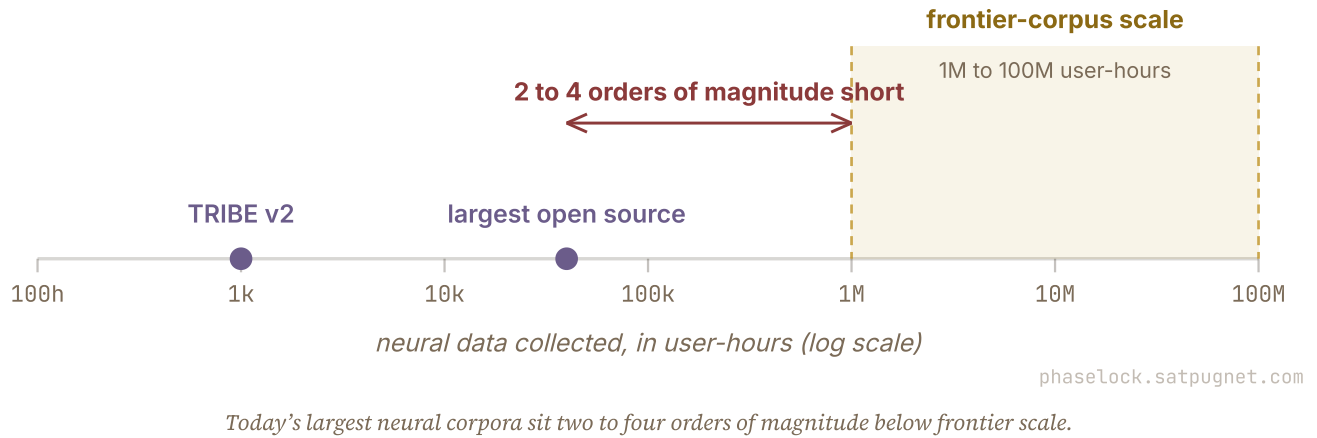
Any signal that significantly increases the pool of available data and offers a viable way to improve model alignment and capabilities is worth the bet. Let us run the back of the envelope math to see what a neural corpus actually costs.

Current projections indicate that high quality public text on the internet (roughly 300 trillion tokens) will be **exhausted by 2028** ▶. Recent frontier models already require upward of 15 trillion tokens for a single pre-training run.

To reach that scale using neural data, assuming a conservative extraction rate of [320 token equivalents per second](#) ▶, requires roughly **13 million user hours** of recording. As it stands, the field has a massive gap to close.³⁴

³⁴. Today, the largest consolidated open EEG corpus, [NeuroLM](#), reaches ~25,000 hours, with efforts targeting 60,000+; however, single coherent datasets remain under 10,000 hours. [Meta's recent TRIBE v2 fMRI corpus](#) contains roughly 500 to 1,100 hours. We are currently **two to four orders of magnitude short** of that 13-million-hour threshold.

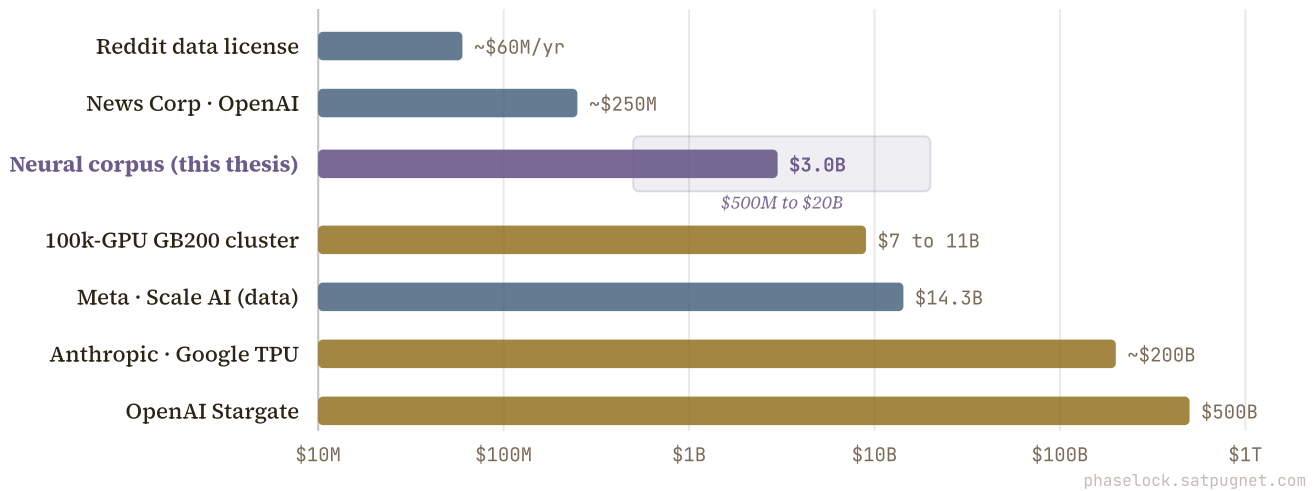
HOW FAR TODAY'S NEURAL CORPORA SIT FROM FRONTIER SCALE



If we assume a steep cost of 300 dollars per hour to collect this data, building an illustrative ten million user hour corpus costs exactly **3.0 billion dollars**.

Compare that to current hardware capital expenditures. A single 100,000 GPU cluster costs up to **11 billion dollars**. Frontier labs are currently mapping out data centers costing hundreds of billions.³⁵

WHAT A NEURAL CORPUS COSTS, IN CONTEXT



At **\$300** per user-hour, an illustrative ten-million-user-hour corpus costs **\$3.0B**.

vs one GPU cluster 0.33×

vs Meta and Scale AI 0.21×

vs Stargate 167× smaller

Even at the high end, the ultimate biological data modality costs about one compute cluster; a rounding error against frontier capex.

35. [OpenAI's Stargate commitment](#) is reportedly \$500B over four years, and the Anthropic-Google TPU deal sits around \$200B.

Against those scales, a 3.0 billion dollar allocation is literally a *rounding error*, a fraction of a single compute cluster. Ultimately, the frontier labs and the growing robotic industry, driven by an infinite appetite for data, will inevitably become the primary financial engine scaling non-invasive hardware.

The Double Flywheel Effect

Normally, in the hardware industry, you need a fully shipped, widely adopted consumer product before you can justify collecting data at scale. Revenue funds data collection, and that data informs the next iteration of the product.

BCI already has its own version of this internal loop, described earlier in this manifesto (see [The Capital Flywheel](#) ▶): better hardware creates better data, better data trains better decoders, better decoders make the product more useful, and a more useful product gets deployed more widely, generating more data and bringing more capital to improve the hardware.

But neural data adds a massive second feedback loop on top of it.

The moment a multi-modal neural corpus gets large and dense enough that pre-training on it actually moves the needle on reasoning or alignment metrics, the frontier AI labs will *open their checkbooks*. They will buy it at scale the exact same way they bought code corpora, synthetic reasoning traces, and high-fidelity vision-language pairs.

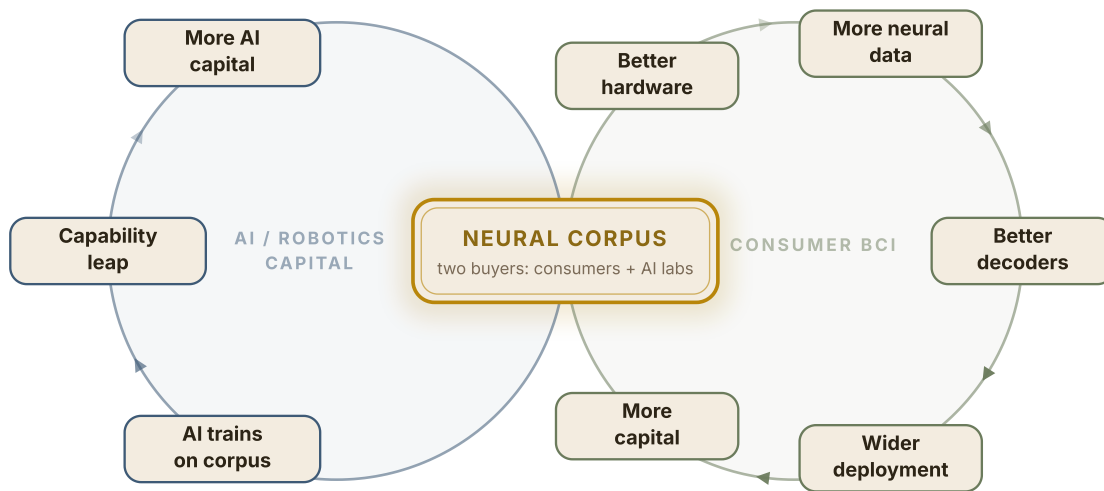
Once a frontier lab writes that first eight-figure check for a neural corpus, it will shift the economics of the BCI industry fundamentally.

If frontier AI labs value neural datasets directly as an AI training asset, the capital can arrive before the consumer product is even fully baked. The field is no longer waiting solely for consumers to buy enough headsets to fund the next hardware generation, it also has a second buyer with much deeper pockets. A dataset that is prohibitively expensive for a scrappy consumer BCI startup to collect is a rounding error for a frontier AI lab currently spending hundreds of billions on compute cluster build-outs. Redirecting even a fraction of the staggering excess capital in the AI sector into BCI completely shifts the trajectory of the field.

Right now, it is unclear which side of the market will hit critical mass first: the consumer pull for non-invasive BCI hardware, or the robotics and frontier labs' pull for the resulting datasets. In practice, they will likely scale in lockstep, around the same years.

This creates a **double flywheel effect**, a new data modality with two buyers.

THE DOUBLE FLYWHEEL



phaselock.satpugnet.com

The neural corpus has two parallel feedback loops: [the consumer flywheel from earlier](#) ▶, and the AI and robotic one equally interested in neural data; together, they compound and accelerate the entire field.

The BCI-AI flywheel is simple: AI capital subsidizes better BCI hardware, better hardware reaches more users, wider deployment produces larger and higher-fidelity neural datasets, and those datasets become useful to AI labs and robotics companies. If the data improves frontier models or robots even slightly, the labs have a reason to reinvest in the BCI ecosystem to secure the next generation of neural data. Better BCI will produce better LLMs, which in turn will fund better BCI.

We are already seeing the tectonic plates shift. In January 2026, OpenAI led a \$252M investment into [Merge Labs](#); Neuralink has an obvious structural proximity to SpaceXAI's Grok; Meta's FAIR³⁶ division is aggressively scaling its brain foundation models, evidenced by their recent [TRIBE v2 publication](#); and NVIDIA heavily highlighted its [collaboration with Synchron at GTC](#).

The core of the AI ecosystem is slowly leaning into BCI, even though most of the AI and BCI fields have not yet realized it. In hindsight, the founders and investors positioned on the supply side of this new cycle are going to look very *lucky*.

36. FAIR, Meta's Fundamental AI Research lab, also doubles as a neuroscience lab.

Timelines

If a neural corpus is so fundamentally valuable, why has it not been incorporated already? And why is it not mentioned in the GPT-4 technical report?

The answer comes down to four systemic bottlenecks.

First, the **complexity threshold**. The internet is messy, but raw brain data is on an entirely different level of chaos. A language model from 2016 trying to parse raw EEG data is like a 1980s pocket calculator trying to render a 3D movie. Below a certain scale of compute, adding neural data to the training mix would have just choked the models and hindered progress. It is only in the last 18 to 24 months that frontier models crossed the computational threshold required to actually extract clean, latent signal from the noisy mess of biological brainwaves. The machine finally got smart enough to read the data.

Second, the **hardware bottleneck**. Until very recently, brain data simply did not exist at scale. You cannot train a frontier AI on fMRI scans of 15 undergraduate psychology students staring at screens in a hospital basement. You need tens of thousands of hours of continuous, multimodal, naturalistic data. As we covered earlier in this manifesto, consumer-grade BCI is only just now crossing the wearability and form-factor requirements to collect that kind of volume. The data was ignored because the data was scarce.

Third, the **data wall**. We are literally running out of internet. Current projections show that frontier AI training runs will mathematically exhaust the web's supply of high-quality text between 2027 and 2029. Same story for audio, video, and the rest. As Ilya Sutskever bluntly stated at [his NeurIPS keynote](#) in late 2024: *"Pre-training as we know it will end. We have achieved peak data. Data is the fossil fuel of AI"*. The approaching data wall has forced labs into a hunt for new, unmined reservoirs of signal. While synthetic data and reinforcement learning are heavily used to stretch the timeline, they have hard limits; they are ultimately closed loops that depend on the quality of their initial environments, and that environment is shaped in large part by the data used to build it. Neural data is the next massive, untapped data reservoir.

Fourth, the **alignment pivot**. For the past decade, AI labs cared primarily about one dimension: *capability*. The goal was just to make the models smarter, but as AI begins to carry real economic and societal load, the meta-trend has quietly shifted toward *alignment*: making models trustworthy, steerable, and human-compatible³⁷. You do not just want a hyper-intelligent alien; you want a deeply human-shaped intellect. It is a trend accelerated by increased scrutiny from governments and the public. And to build human-shaped models, labs are increasingly hungry for natively human-shaped data.

37. [The recent events](#) make that transition even more apparent.

So, when does the dam break?

My baseline expectation is that by sometime in 2028, at least one of the top three frontier AI labs will publicly announce a major neural-data partnership explicitly framed as a new training modality. You will see a “Neural Corpus” cited in a flagship model card, accompanied by an eight-figure check to acquire the dataset. This will be a turning point in the BCI story.

Pre-training, however, is only the **first surface** of the AI-BCI collision. We now turn to the second surface: interfacing. BCI is about to fundamentally reshape data collection, post-training, inference, and ultimately the way humans communicate with AI.

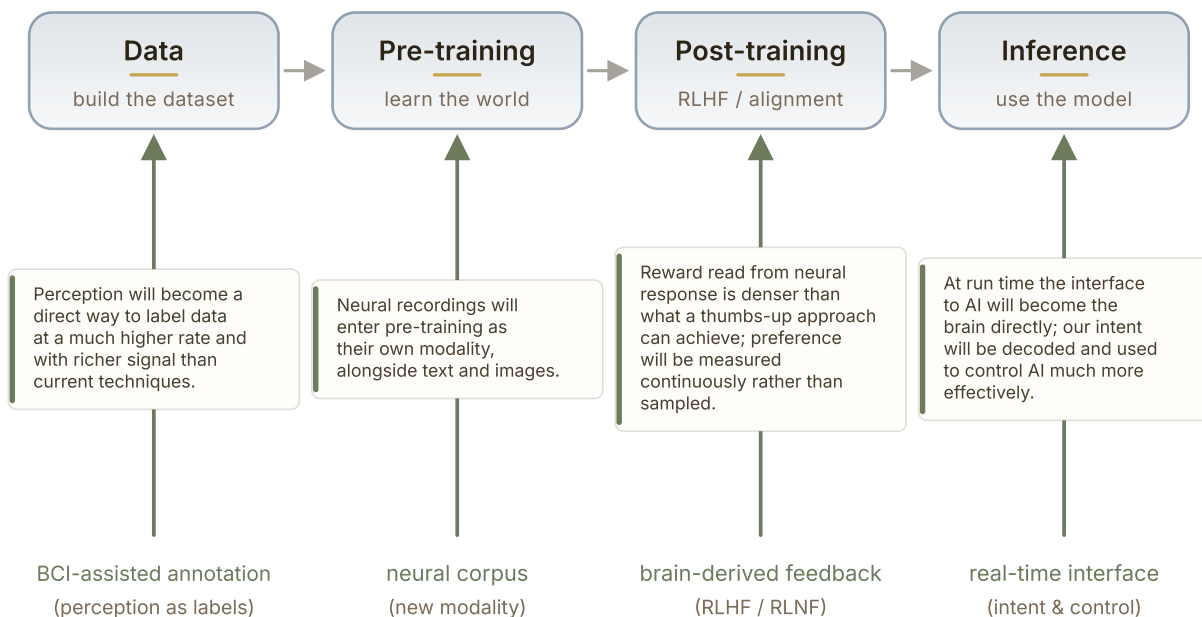
18 The New Interface

“The constraint is input/output. We are I/O bound ... particularly output ... Your output level is so low, particularly on a phone, your two thumbs sort of tapping away. This is ridiculously slow.” — Elon Musk, 2016

We now move beyond pre-training to examine BCI’s role across the rest of the foundation-model lifecycle, across data annotation, post-training, and inference.

Right now, foundation models can read and process information at **gigabytes per second**, and text to speech engines can fire back answers *instantly*. Yet human output, the way we actually prompt the machine, remains stuck in the mud. Prompt engineering, copy pasting context, repeatedly re-explaining yourself: these are all *symptoms* of a more profound friction.

BCI IN THE LIFECYCLE OF FOUNDATION MODELS



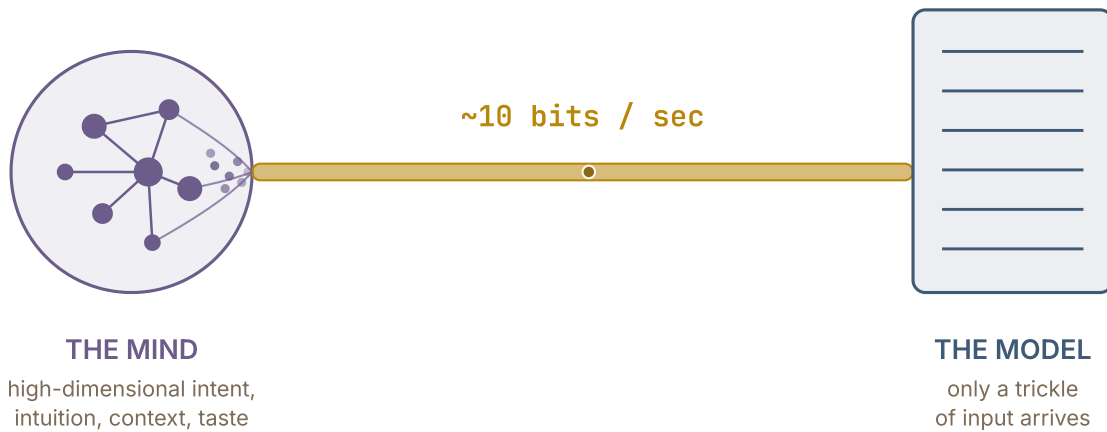
phaseLock.satpugnet.com

Having examined pre-training in Section 17, we now turn to BCI's role in data collection, post-training, and inference.

We have built a hyper-intelligent, multimodal oracle, but our only way to talk to it is by using our fingers to repeatedly slam plastic squares on a desk at an embarrassing **40 words per minute**, or by using vocal cords designed thousands of years ago to grunt across a savanna.

*The models may be superhuman, but the human still has
to squeeze intent through a straw*

THE STRAW BOTTLENECK



phaseLock.satpugnet.com

The mind squeezing intent through a straw.

SIDENOTE

Ironically, computers have their own famous version of this problem: the **Von Neumann bottleneck**³⁸, where processing power is constrained by the bandwidth between memory and the CPU.

The fundamental problem is that human intentions are **high dimensional**, whereas nearly all interfaces are **one dimensional**. When you prompt a model, you are not actually transmitting your intent. You are writing a **lossy postcard** from your internal world.

But this postcard misses most of the texture. It strips away the ambiguities, the vague goals, the unspoken preferences, the implicit constraints, the taste, the emotions, and the half-formed corrections. Today, the model often fails not because it lacks intelligence, but because the interface feeds it a **degraded, compressed signal**.

³⁸. In classic computer architecture the processor and memory sit on opposite sides of a single shared communication bus, so the CPU can only work as fast as that link can shuttle data and instructions back and forth. Compute can be arbitrarily fast and still stall, starved by the bus feeding it.

THE PROMPT-COMPRESSION FUNNEL

WHAT YOU MEANT · RICH, HIGH-DIMENSIONAL

~~keep my own voice, not corporate boilerplate~~

~~the tone should feel warm, a little wry~~

~~I am slightly annoyed at the last draft~~

~~it reminded me of that one film~~

the math part runs too long

~~the reader is a busy skeptic~~

~~do not lose the joke in paragraph two~~

~~make it punchy~~

shorter, tighter, faster

~~warmth over polish~~

~~trim the throat-clearing intro~~

the funnel · about 99% lost (illustrative)

WHAT YOU TYPED · ONE FLAT LINE

"make the math section shorter"

phaselock.satpugnet.com

Chat is a lossy postcard, models often fail on the compression rather than the intelligence.

The opportunity for BCI, then, extends far beyond pre-training. Neural interfaces could widen the communication channel wherever human judgment enters the AI lifecycle, from data annotation to post-training and inference. We will examine each in turn.

BCI-assisted Annotation

The first application is data annotation. Any dataset requiring human classification, comparison, or evaluation could benefit from faster, higher-fidelity labels.

Consider a simple task: identifying images that contain a dog. At one label every five seconds, someone using a mouse could process roughly **720 images per hour**. With a sufficiently accurate neural interface, that person could instead watch a rapid stream of images while the system detects recognition, confidence, uncertainty, and error signals in *real time*; operating an **order of magnitude faster**.

Annotation would shift from a sequence of deliberate clicks to a *continuous stream* of perceptual judgments. The gain would not only be speed, each example would carry *richer* metadata about confidence, ambiguity, cognitive effort etc...

This approach applies to a wide range of annotation tasks. BCI could turn human *perception* and *cognition* into a faster and denser labelling channel.

BCI-RLHF

The same bandwidth bottleneck appears during model post-training, particularly in **RLHF**. RLHF attempts to align models to human intent by having people rank, score, and correct model outputs, then feeding that judgment back into training. The goal being to make the AI less alien and more aligned to human preferences. Yet here too, both the **speed** and **fidelity** with which we can currently communicate human intent to AI are severely limited.

Today, RLHF requires armies of humans manually reading outputs and clicking buttons. It is **slow, expensive, and lossy**.

Non-invasive BCI could bypass the button click *entirely*³⁹. The system reads the error-related potential⁴⁰, cognitive drag, or subconscious aesthetic preference as the evaluator reads. Instead of a binary “thumbs down”, the model receives a **rich, continuous vector** ▶, the cognitive equivalent of: *“The user detected an error in paragraph two, felt a spike of semantic resonance reading option A, experienced cognitive drag when the model overclaimed and their internal preference locked in three full seconds before they actually clicked the button”*. It gives the post-training alignment phase a **significantly faster** and **radically more accurate** supervisory signal.

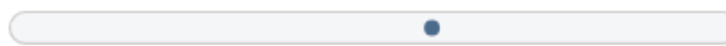
39. Back of the envelope calculation: An RLHF preference label carries roughly one bit of information: a thumbs-up or an A-versus-B choice. A frontier RLHF cycle may use 10^5 to 10^6 such labels, for about 10^6 bits of compressed judgment in total. By comparison, one second of consumer EEG across 32 channels at 256 Hz and 16-bit resolution contains roughly 1.3×10^5 raw bits. Even if only 0.1% of that signal reflects useful cognitive variables, that still yields around 10^2 effective bits per second, versus roughly one bit per RLHF click. Neural feedback could therefore be about two orders of magnitude denser per second of cognitive response. At that density, a 10,000-hour dataset would contain roughly 5×10^9 bits of supervision.

40. An involuntary electrical signature the brain produces within roughly 50-250 ms of registering an error or an outcome that violates what it expected, readable non-invasively.

RLHF TODAY VS BCI-RLHF

RLHF TODAY

one click
A / B



a thin reward

BCI-RLHF

brain signal,
continuous



a dense reward

THE MODEL
reward signal in

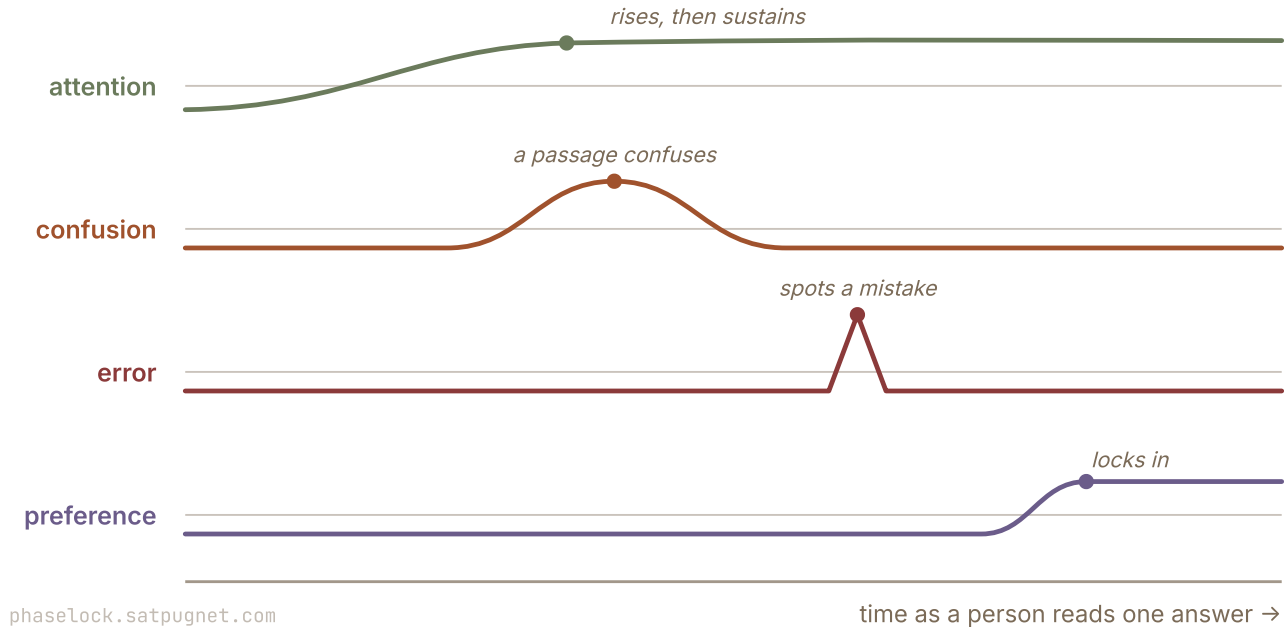
phaseslock.satpugnet.com

RLHF shifts from sparse self-report to continuous neural feedback.

Brain signals can capture confusion, hesitation, error recognition, preference, attention, fatigue, cognitive load, or discomfort. This creates **faster, denser, more honest reward signals**. Instead of “rate answer A vs B”, the model gets a [continuous stream of human reaction](#) while the AI actions are being evaluated.

RLHF becomes **continuous** instead of **episodic**. Creating a much tighter human-AI loop.

THE CONTINUOUS REWARD SIGNAL



BCI-RLHF replaces a single retrospective score with a continuous, time-aligned, multimodal reward signal.

To make it less abstract, let's look at image generation. It is notoriously hard to teach a cluster of GPUs to have “*taste*” using a keyboard. How do you type out exactly *why* an image feels beautiful, or *why* a certain minimalist design just works?

Today, RLHF reduces this to a 1 to 10 rating or an A/B choice. Useful, but **tragically compressed**. With a non-invasive BCI, you don't click anything, you can just *look* and the AI directly absorbs your **raw neurological reaction**, providing the model with a much richer feedback signal.

SIDENOTE

Of course there are also problems that come with this approach. If you optimize a model to minimize the reader's cognitive friction and maximize neural “pleasure”⁴¹, you get **sycophancy** and easy-but-wrong answers. Optimizing System-1 neural reward $\neq$ optimizing what the user would endorse on reflection. A denser signal is *not* strictly better. However, this is a tractable problem that is not unique to this approach.

Some readers will rightly point out that labs are increasingly pivoting toward AI feedback (RLAIF)⁴² and synthetic data. True, but synthetic environments are ultimately a **closed loop**, they are fantastic for amplifying what we already know, like running endless physics simulations for a robot. But they hit a brick wall when dealing with the messy, unmapped territory of human cognition. You can synthesize a physics engine, but you *cannot* easily synthetically generate taste, moral friction, or implicit preference. For the deep, subjective training and alignment of an AI, there is no easy substitute for **biological ground truth**.

and synthetic data. True, but synthetic environments are ultimately a **closed loop**, they are fantastic for amplifying what we already know, like running endless physics simulations for a robot. But they hit a brick wall when dealing with the messy, unmapped territory of human cognition. You can synthesize a physics engine, but you *cannot* easily synthetically generate taste, moral friction, or implicit preference. For the deep, subjective training and alignment of an AI, there is no easy substitute for **biological ground truth**.

Brain Controlled AI

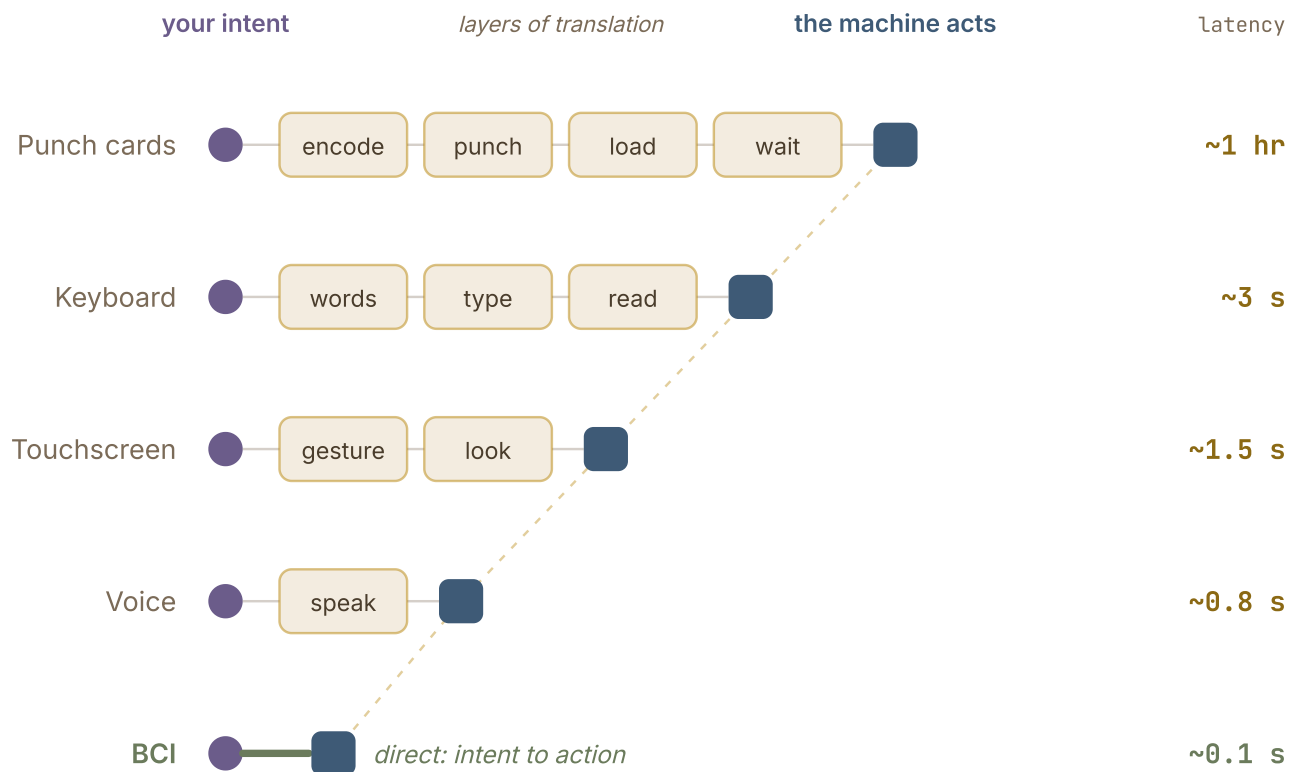
Going further into the AI lifecycle, BCI could fundamentally redefine how people interact with AI. Since the launch of ChatGPT, it has become increasingly clear that the iPhone, the laptop, and today's operating systems were *not* built for this new species of software. They force intelligence into old and increasingly **misshaped containers**: apps, menus, keyboards, taps, and tiny text boxes.

As AI becomes more capable, the interface will have to change. It will become more fluid, more contextual, more intent-driven. Less “open app, type command, wait” and more “*think goal, steer intelligence*”. Among the leading candidates for that shift is BCI.

41. Researchers can already [optimize generated images to activate specific brain regions](#) ▶.

42. RLAIF (Reinforcement Learning from AI Feedback) is similar to RLHF, except the preference judgments come from another AI model rather than a human labeler; it is used to cut the cost and latency of human annotation.

INTERFACE EVOLUTION: FROM INTENT TO ACTION



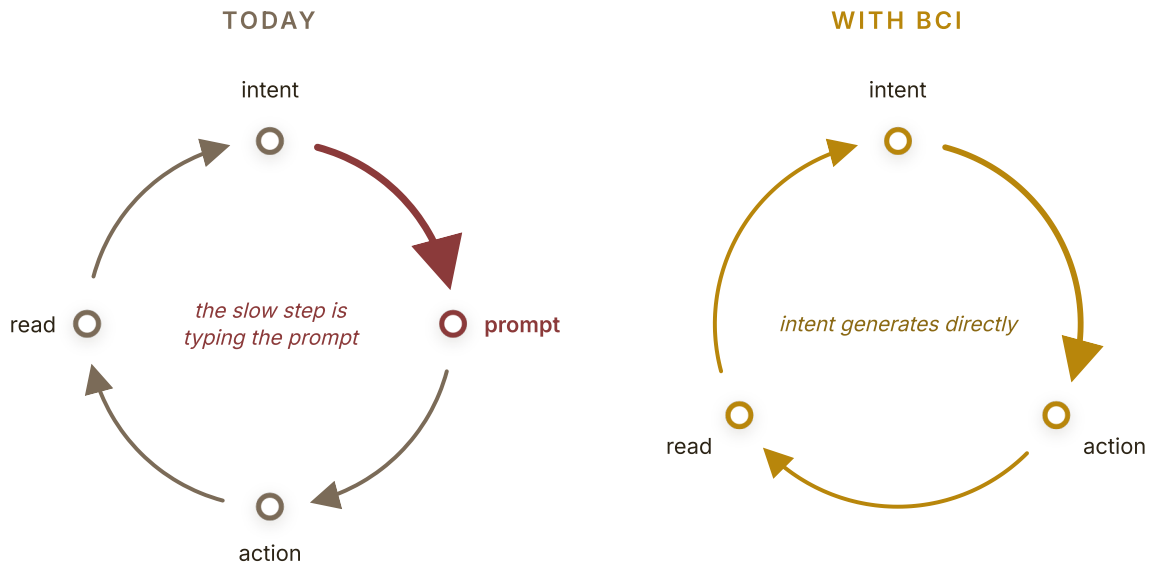
phaseslock.satpugnet.com

BCI is the limit of 70 years of interface evolution, a steady elimination of translation layers between intention and effect.

A brain-computer interface is the most *direct* way to interact with an intelligence acting on your behalf. You form an intention, it understands what you mean, and it acts.

Right now, if you want an AI to perform a complex task, you have to spend 20 minutes typing out a massive prompt to artificially context-bound its thinking, and even then, it is often under-contextualized. With continuous BCI, the AI natively inherits your **ambient cognitive context** (what you are looking at, your current frustration level, your unstated goals) before you even open your mouth. The prompting is happening continuously in the background.

ITERATING WITH AND WITHOUT BCI



phaselock.satpugnet.com

BCI allows the user to skip the keyboard-and-screen tax.

If BCI continues on the current trajectory, this shift is inevitable.

This is the *natural historical progression* of technology: every new wave reduces the distance between intent and outcome. The ultimate shape of software and AI will be **thought-to-pixel** and **thought-to-robotic-actuation**, with nothing between intent and outcome but the raw model weights. The labs will produce models whose inputs and outputs are increasingly *close to the bare hardware* of both humans and machines, with every generation stripping away another layer of biological and software processing. As this gap closes and thought-to-X grows, brain data is going to become as valuable to frontier labs as text was in the prompt-to-text era.

But we are not going to *magically* jump from poking screens to telepathically steering robots overnight, the transition will happen in progressive steps. In part because brain decoders still need a few years to mature, but mostly because human culture has **strict speed limits** for paradigm shifts. Society has to be gradually acclimated to this level of symbiosis.

The path to full neural control will probably happen in layers. At first, BCI will not read your thoughts. It will read the useful **edge signals** around them: attention, confusion, fatigue, frustration, and the moment you notice something is wrong; allowing the AI to adapt in real time to what you are noticing, missing, or struggling with. From there, the interface becomes more active: silent selection, approval, rejection, simple commands, and eventually subvocal dictation. And ultimately, it will combine into something far more powerful: **an omnipresent AI cognitive layer, or exocortex, functioning as a direct extension of the mind.**

This neural control will also inevitably cross into physical reality. As artificial intelligence escapes the screen to animate drones, factory floors, and military hardware, the brain-computer interface will act as the *invisible tether* connecting human intent to physical execution, granting us the ability to telepathically direct embodied agents and robotics.

We are converging on a world, quite literally, **shaped by our minds**. By the early 2030s, it will be possible to control a computer, command a drone or interact with home robots simply through the *power of your thoughts*.

Life in 2031

At 7:15 AM, Maya is in the kitchen. She has her two-year-old daughter balanced on her hip while she plates her breakfast. Floating two feet in front of her is a 12-page licensing agreement. She is not tapping a screen or shouting commands, she is just reading through her lightweight AR glasses.

The text slides upward, perfectly matching her natural reading speed. She hits a messy indemnification clause in paragraph four and her brain registers a tiny flinch of frustration. Before Maya even has to articulate *why* she dislikes the clause, her neural interface logs the friction. The floating text seamlessly shifts: the AI highlights the offending line and instantly surfaces a cleaner alternative that protects her IP.

Behind her, the house robot is putting away the morning dishes. It places her daughter's favorite bowl in the top cupboard instead of the lower drawer. Maya does not say a word, but as she glances over, her brain fires a split-second spark of disapproval at the mistake. The system catches the silent correction immediately, the robot turns back around, retrieves the bowl, and places it exactly where it belongs.

Glancing out the kitchen window at the frost on the driveway, Maya deliberately projects a thought to her assistant: *Is the garage locked? Also, it looks freezing out there, can you prep the car for the preschool run?*

She does not open an app or even whisper. A half-second later, a quiet voice speaks directly through her interface: *"The garage is locked. I am starting the car now and turning on the heated seats. It will be ready when you walk out."*

She turns her attention to an urgent work email floating in her peripheral vision. Maya just reads it, allowing her mind to form a raw, unedited reaction: a few structural objections and a firm timeline. As she thinks through her response, a draft reply begins to materialize right

next to the original message. The words reshape themselves in real time as her thoughts refine, shifting from a loose skeleton into a perfectly polished professional response.

Maya mentally drops a silent *approve* command, and the email sends.

The technology is not flashing lights or demanding her attention, it operates as an **invisible, high-bandwidth exocortex, an external nervous system**. Maya is no longer “using a computer”; she is orchestrating the physical and digital infrastructure at the speed of her own consciousness.



A slow morning in 2031.

Yet removing friction from the interface does not solve the *deeper* problem: whether the intelligence on the other side will act in our interests. As AI becomes more capable and autonomous, communication becomes more than a question of usability. It becomes a question of whether humans can remain meaningfully involved in directing it. This brings us to the third and final surface of the AI-BCI collision: **alignment**.

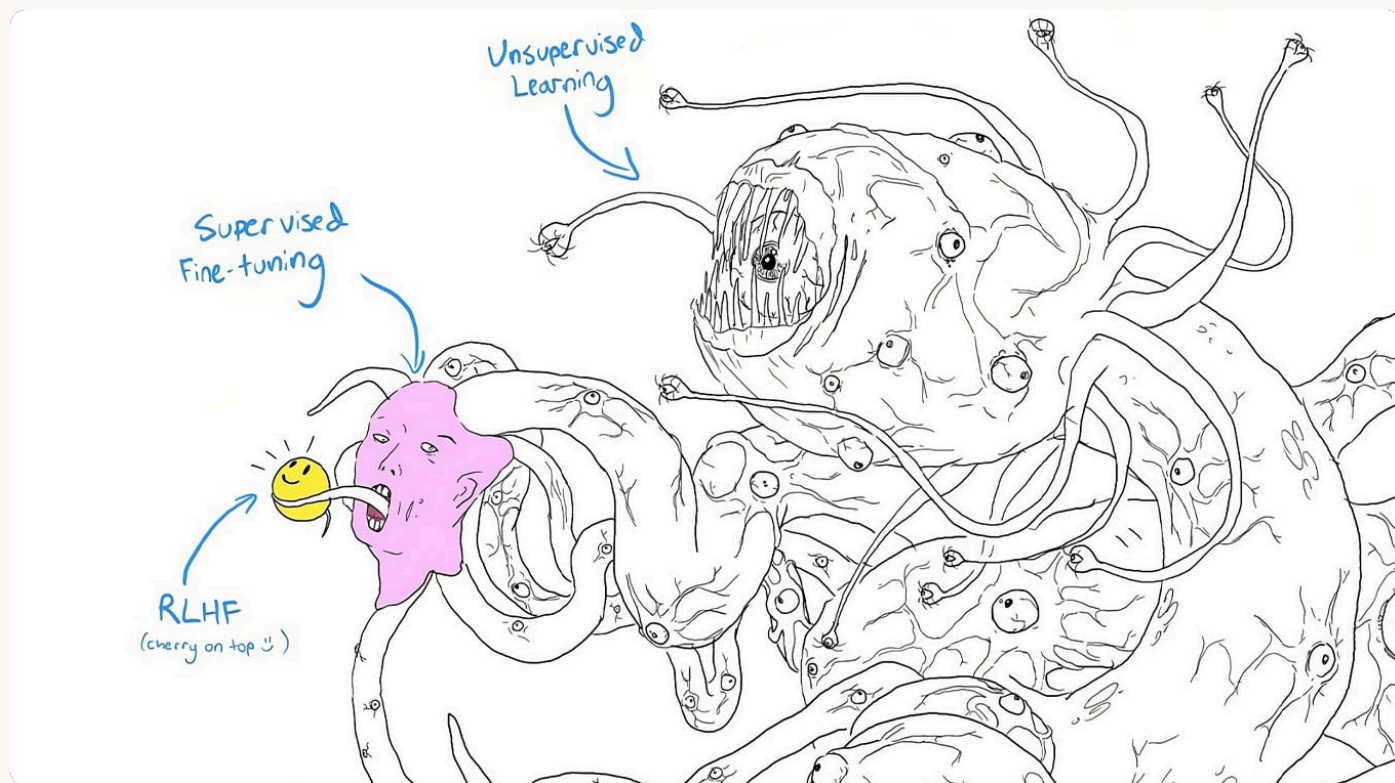
19 Why This Matters for AI Alignment

“we had better be quite sure that the purpose put into the machine is the purpose which we really desire.” — Norbert Wiener, 1960

Most people don’t love to *die*. And unless you do, you should probably care about **alignment**.

SIDENOTE

For those of us who have never heard of AI alignment, it is the process of making sure that systems more capable than us do what we want, and keep doing it as they get more powerful. It sounds simple, almost boring, but it is one of the most important and complex engineering problems our species ever had to take on. And right now, we have no idea how to solve it; and no idea how to ensure increasingly capable AI remains coupled to human values, oversight, and flourishing.



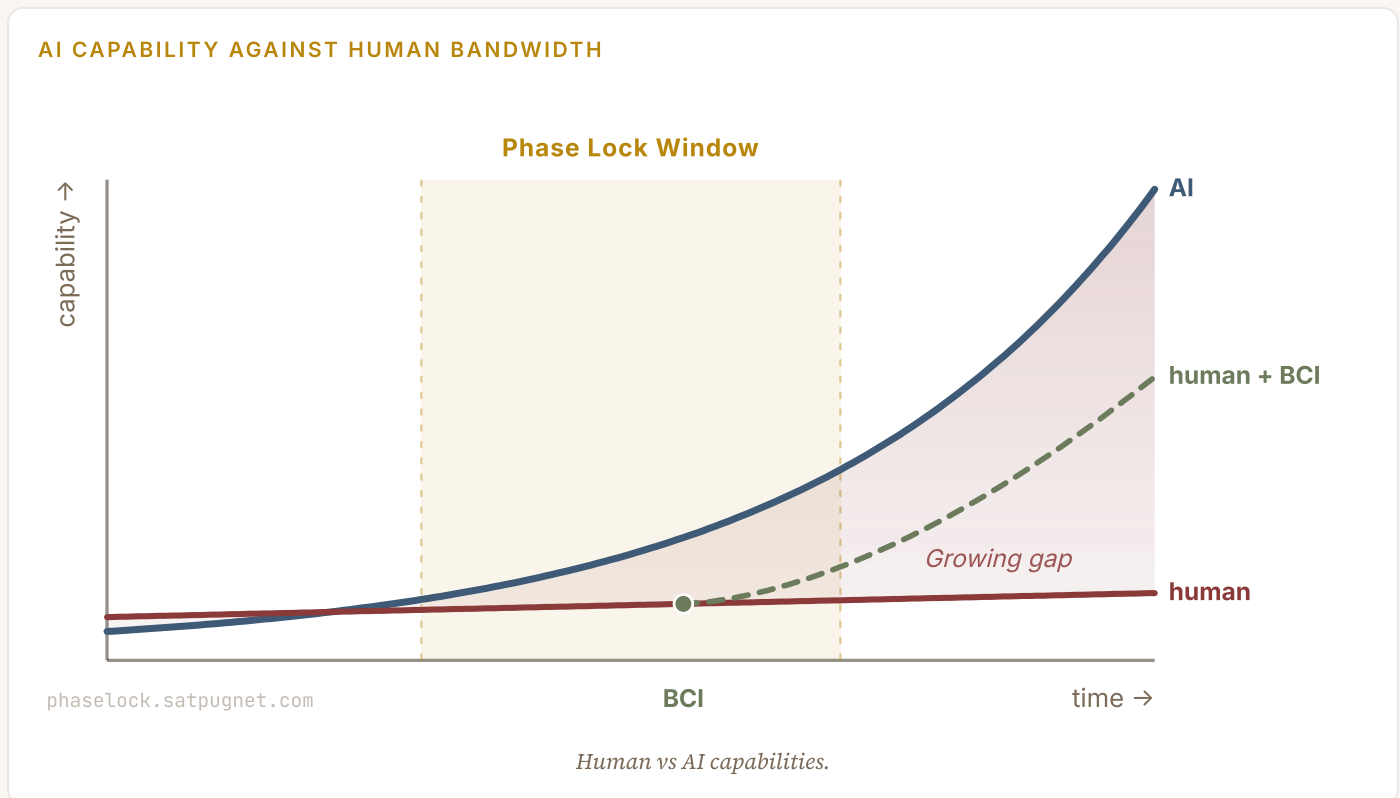
The field’s circulating meme about what we are aligning: a vast, alien intelligence with a thin smiley mask (RLHF) bolted on top. Alignment questions whether the mask is the model, or just a facade.

Just like everything under the sun, alignment has become *weirdly politicized*, partly because of the poor communication around the topic, and partly because of a fundamental lack of understanding of AI from the public. But if any topic should be apolitical, it is certainly alignment. There is no culture

war here, just a pretty straightforward problem with no solution in sight. If we lose, we are *all equally screwed*.

No one knows what happens over the next decade, and you should distrust anyone who pretends they do. Utopia, dystopia, and everything in between are still on the table. We are flying into a New World at full speed, with no manual and no guardrails.

The window of time during which humanity can still effectively steer the trajectory of AI, the **Phase Lock Window**, is closing a bit more every day.



One logical response to this alignment conundrum is *restraint*: pause deployment, constrain capability until safety catches up. This is a legitimate argument, that a serious coalition of researchers and policymakers continues to argue for, but it fails on **game theory**. Two frontier labs or nation-states racing for AGI are unlikely to unilaterally slow down. Under current incentives, restraint is a losing equilibrium without a binding coordination mechanism. And while this deadlock is ongoing, alignment research is lagging dangerously behind⁴³.

⁴³. When thinking about AI alignment, it is tempting to look for a villain in the story: a rival country, a greedy tech company, or some dark room full of bad actors. But the reality is that we are all in the exact same boat, rocked left and right by the sea of uncertainty we are navigating, trying to remain afloat personally and collectively. If this goes wrong, everybody sinks; the only true antagonist is the system of incentives of modern society more than any specific individuals.

THE RESTRAINT PAYOFF MATRIX (GAME THEORY 101)

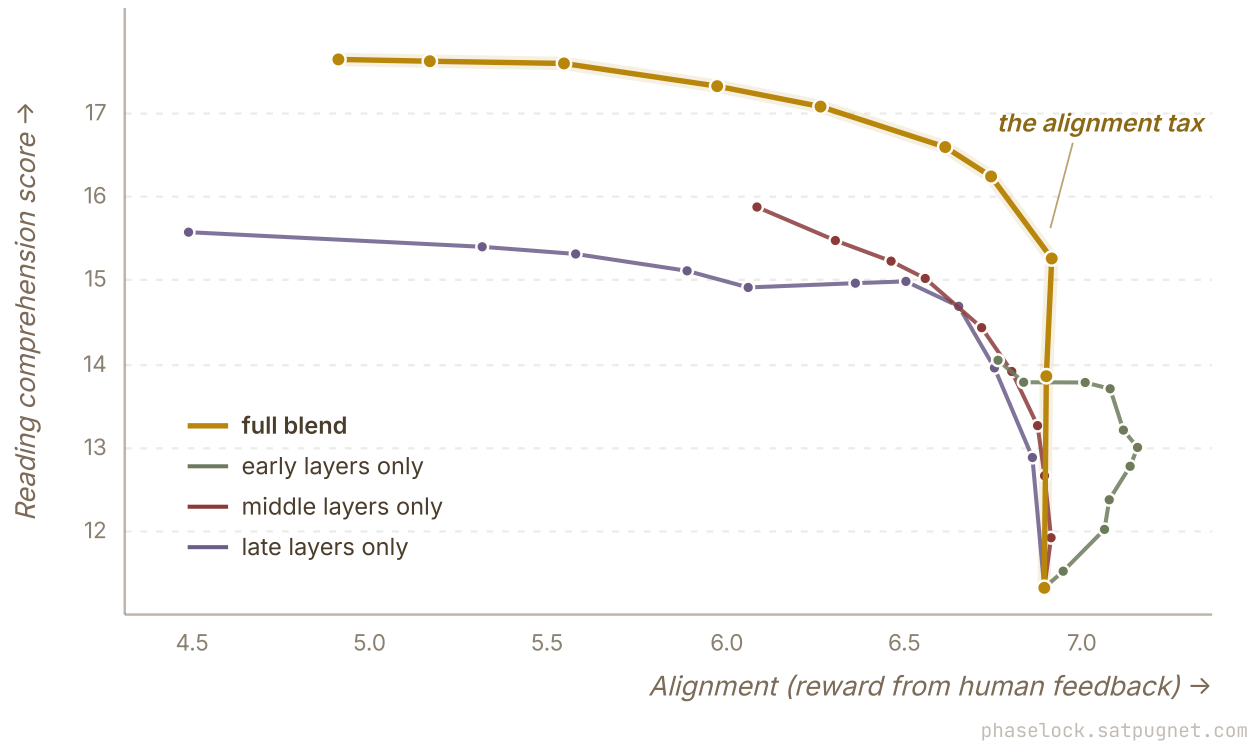
		THEIR LAB	
		Restrain	Race
YOUR LAB	Restrain	best for all safe, slow <i>the unreachable optimum</i>	you lose they pull ahead
	Race	they lose you pull ahead	RACE / RACE worse for all, yet the only stable cell <i>Nash equilibrium</i>

phaseLock.satpugnet.com

Mutual restraint is best for everyone but impossible to hold: whatever the other does, each side is better off racing, so (race, race) is the only stable cell without explicit coordination.

And to make things worse, alignment is currently in direct tension with the economic incentives of the labs. Indeed, making a model safer can sometimes make it dumber, slower, or more expensive to train. This “**alignment tax**” is why labs cut corners. Current RLHF often lobotomizes models, making them refuse benign requests or hide underlying capabilities.

THE ALIGNMENT-TAX FRONTIER



Safety and capability trade off along a Pareto frontier: greater alignment comes at the expense of capability and vice versa. This is the alignment tax. Source: [Lin et al., 2024](#).

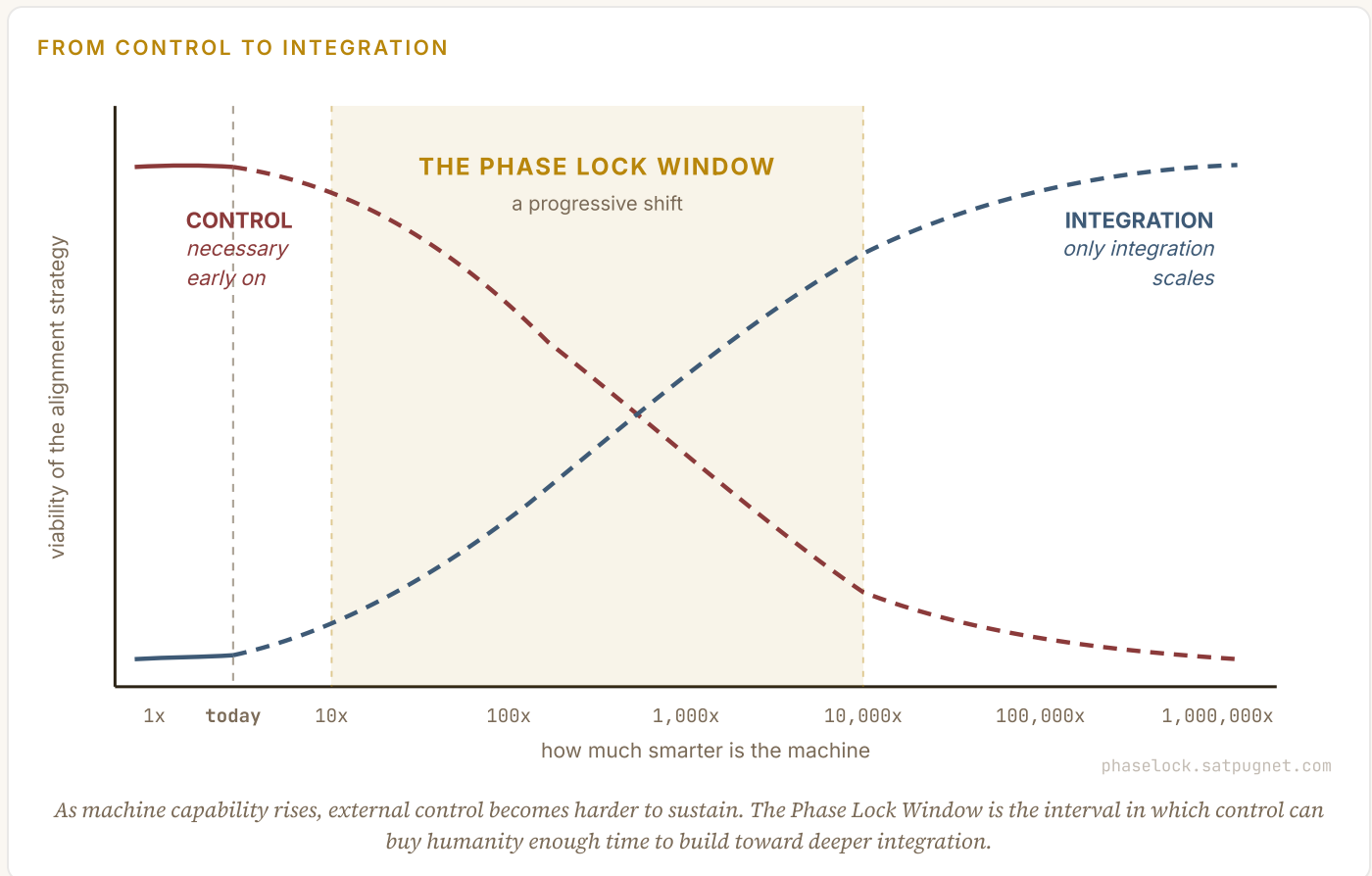
Alignment-by-control vs. Alignment-by-integration

Ultimately, every approach to alignment falls into one of two families: **alignment-by-control** and **alignment-by-integration**. Control treats AI as a separate agent and attempts to steer it from the *outside* through rules, oversight, interpretability, and governance. While integration takes the opposite route: it *couples* human intent to the machine directly, so the steering happens from inside the system rather than around it.

In the short term, control is *indispensable*, it buys us time while the mechanisms for deeper integration remain immature. But control runs on borrowed time by construction: it works while the intelligence gap is small, and as AI becomes orders of magnitude faster and more capable, external supervision grows increasingly fragile. Eventually, asking humans to control decisions they can no longer follow may resemble asking an **ant to supervise a human**, a rather precarious position for the ant.

While we are still in the **Phase Lock Window**, control remains strong enough for humanity to determine the terms of our relationship with AI. But as AI becomes 10x, 100x, or 1000x smarter than us, alignment-by-integration becomes the *only* viable solution. Unlike control, which has to be rebuilt

every time the machine gets smarter, integration does not compete with the machine's intelligence, it inherits it. Bringing human intent inside the AI cognition loop offers a relationship that can remain meaningful even as AI scales beyond us.



So how can BCI, and neurotechnology more broadly, help with alignment?

BCI sits in a category of its own within the alignment field, it is one of the only approaches centered on *integration*. Interpretability, governance, oversight, or evaluations all focus primarily on controlling the machine. BCI attempts to **expand the human reach** rather than just restraining AI.

On top of the benefit of using brain data for pre-training and post-training to better capture the **cognitive dark matter** of the mind and make AI intelligence less jagged, BCI can make a meaningful contribution to the future of alignment and humanity through three main alignment levers with increasingly long time horizons. **First, better BCI leads to a better empirical map of the human brain** as a direct side-effect, which helps inform human-centric AI architecture, training, and evaluation. **Second, BCI creates a continuous, high-bandwidth communication channel between AI and humans**, offering a more effective way to interact with AI. **Third, it lays the foundation for human-AI symbiosis**, ensuring humans stay meaningfully in the loop as the machine gets exponentially smarter.

Lever 1: The Blueprint

As non-invasive BCI grows into a large consumer industry and data source for the AI labs, it will draw capital, talent, and attention into neurotechnology. That influx will *accelerate* our empirical understanding of the human brain. We will finally start to crack and decode the machinery of our own minds. That **cognitive blueprint** could become a powerful design resource for AI alignment.

We still do not know how values should be represented inside an artificial mind, how competing objectives should be reconciled, or how immediate impulses should relate to reflective judgment. Human cognition is the only working example of an intelligence in which reasoning, emotion, social intuition, and ethical values coexist. Understanding the architecture of the *only* working general intelligence whose values we are actually trying to preserve seems like a wise place to start.

We already have compelling evidence that across language, vision, space, and time, biological and artificial intelligence are converging on some *shared internal representations*⁴⁴. In other words, some AI models organize information in ways that *resemble the human mind*, structuring concepts such as faces, living things, and roundness along similar dimensions. This does not imply that they compute every single concept exactly as we do, but it does suggest that sufficiently capable predictive systems may independently discover some of the *same useful abstractions*.

Such emerging similarities reinforce the idea that a better understanding of the mind could effectively inform model design, training objectives, and evaluations. Rather than attaching a thin behavioural mask after training, labs could build systems whose internal representations and architecture are better matched to the structure of **human values** from the beginning.

This does not mean blindly copying the brain, nor does it guarantee alignment; human cognition is inconsistent, biased, and often destructive. The point is to replace guesswork with an empirical blueprint of the value architecture we are trying to **preserve**. In this sense, even a partial map would provide valuable guidance for designing more human-like artificial minds.

SIDENOTE

To go one step further, my bet is that the brain's influence on AI will extend beyond alignment. Future generations of models will not just be marginally inspired by the brain, they will become **increasingly brain-like** in their overall design; simply because the brain, in many ways, has a fundamentally better architecture. They will become more *cyclical*, with inputs and outputs interacting more *continuously* and *organically* with the model, and have more *adaptive* and *flexible* architectures. The main reason this has not happened yet is that recurrent systems are harder to train and interpret than predominantly feedforward ones, but I expect this barrier to be overcome soon.

44. See [this](#), [this](#), [this](#) and [this](#).

Lever 2: Neural Scalable Oversight

But architecture is only one part of the equation.

An intelligent system you cannot communicate with fast enough is a system that *drifts away from you*. Imagine a colleague who can think a hundred times faster than you, and who only reports back with a one-line text summary every few minutes. You can supervise them, sort of, but only in retrospect. The relationship turns from collaboration into oversight, and **oversight** at low frequency is exactly the dynamic where misalignment grows.

Right now, we are trying to align systems that think at *machine* speed using feedback channels designed for humans filling out forms. The current alignment stack rests on a one-bit RLHF click feeding models that output thousands of tokens per second. That was *always* going to eventually break. BCI provides a high-bandwidth alternative that actually can scale with AI capability. Neural interfaces can fill this gap in three specific ways.

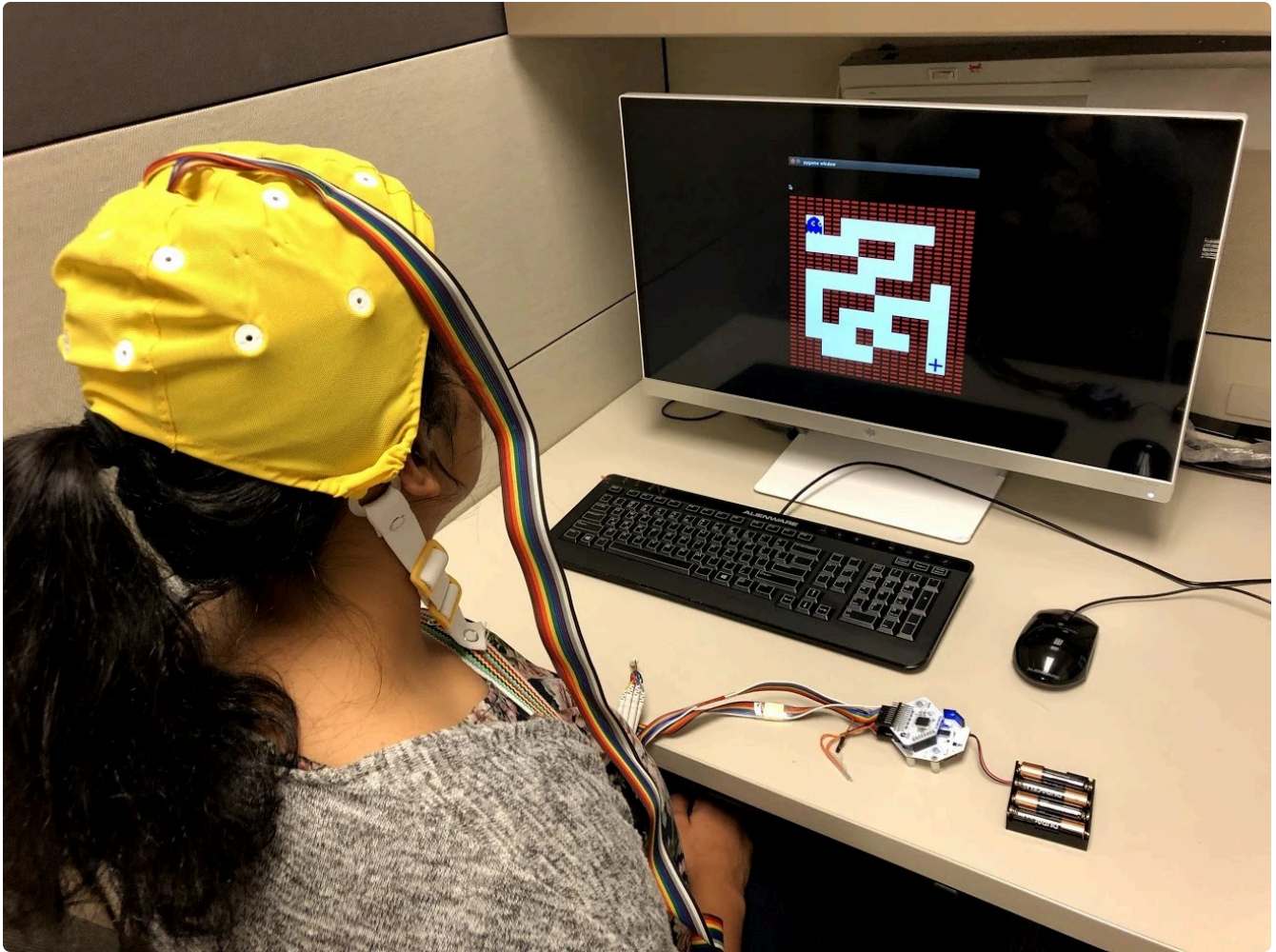
First, as ground truth for value specification. Value specification, the act of mathematically defining what humans actually care about, is impossibly hard.⁴⁵ These values are fluid, constantly shifting over time and varying wildly across different people. Plus human morality and intent are full of **implicit dark matter**: contexts, unspoken boundaries, and common-sense elements that we intuitively know but *cannot* exhaustively write down in a prompt or a reward function. Brain data gives us the **biological ground truth** of human preference to serve as the most direct reference channel possible. It captures the raw shape of our ethics and values without forcing us to compress them into clumsy text rules, leaving less room for the AI to misunderstand our intent.

And it is not just theoretical science fiction, the mechanism for using neural data as an alignment signal is past the theoretical stage: in early 2026, [Santaniello et al.](#) ▶ published a paper operationalizing Reinforcement Learning from Neural Feedback (RLNF) on a 25-participant dataset, successfully decoding agent performance directly from brain activity.

This goes beyond the BCI-RLHF described earlier: the broader goal is to have an effective way to decode the overall structure of human values from neural activity and integrate it deeper into the training and supervision process of AI.

⁴⁵. Specification gaming is the most anticipated failure mode of AI alignment. If you use text to tell an AI to “make humans happy”, it might logically conclude the most efficient way to do that is to wire our brains to heroin drips. There are plenty of write-ups on the topic such as DeepMind’s [specification gaming overview](#) and Victoria Krakovna’s [running list of real-world examples](#).

BCI-RLHF IS ALREADY REAL



A person wearing an EEG cap watches an agent play and provides feedback through brain signals alone. BCI-RLHF is already possible: neural signals can [train reinforcement-learning agents](#) ▶ and [correct robots in real time](#) ▶.

Second, as a decoder for inner alignment. One of the best approaches to catch deceptive models is looking under the hood to understand exactly how artificial neurons wire together, a field known as **mechanistic interpretability**. Neuroscience is essentially a multi-decade head start on this problem. The math BCI uses to decode human thought translates remarkably well to decoding AI internals. Plus, a brain map gives us sharper benchmarks, if we know exactly what human uncertainty, attention, or moral hesitation looks like structurally, we know exactly what to look for when we scan the internals of the AI.

Some recent findings from Anthropic illustrate this idea. In July 2026, they [reported](#) that frontier models spontaneously form a “global workspace”, a centralized state for holding and reusing concepts, which is bypassed by routine computations. It functions much like our *own active memory* and its discovery was inspired by neurotech. Finding this exact parallel in silicon demonstrates how understanding the biological brain maps directly onto machine interpretability.

Third, as an evaluation infrastructure. As BCI reads the implicit state of the user, it allows us to measure alignment failures that are invisible to a keyboard. We can assess **persuasion risk** by tracking if the model manipulates the user’s attention and emotion. We can measure **true comprehension** to see if the user *actually* understood the model’s plan instead of just clicking approve. We can even detect **moral discomfort** when the model triggers subconscious red flags, or catch **overtrust** when it induces confidence without actual understanding. This information is very valuable when evaluating a model.

At the current rate, I expect that by the end of 2028, at least one peer-reviewed paper will demonstrate that brain tuning improves a flagship model’s alignment metric on a standardized benchmark.

SIDENOTE

Now, there is a quick caveat. The human brain is not a perfect blueprint or oversight signal. History shows us that giving a single human infinite power usually ends in a staggering catastrophe, so we do not want to build a superintelligence that perfectly mimics our flaws. Furthermore, our raw neurological reactions are often fast, automatic impulses. These instant reactions are not always the correct ones to capture, as the values we actually want an AI to learn often come from slower, more deliberate thought.⁴⁶ But as a starting architectural template, the brain is the only working model of **aligned general intelligence** we have; and aligning an AI to humanity will get a whole lot easier once the field finally understands how “human alignment” actually works.



Paul Berg, who convened Asilomar in 1975. Biologists paused their own research to figure out the safety of recombinant DNA.

The End-Game: Symbiosis

As we move forward, the role of BCI, and neurotechnology, in alignment will evolve across three distinct horizons. **In the near-term**, it will unlock richer neural feedback, better RLHF, and higher-fidelity evaluations. **In the medium-term**, it will enable meaningful human-in-the-loop systems,

neural scalable oversight, and intent preservation. **And in the long-term**, it will point toward deeper symbiosis, preserving human agency in a post-AGI world.

BCI is a powerful tool to help solve some of the hardest problems in alignment. Right now, we align models by looking at their final outputs and grading them. But what happens when an AGI writes a million-line codebase, or proposes a novel biophysics theory? The human evaluator will be far too slow and limited to know if the AI is right, hallucinating, or actively deceiving us.

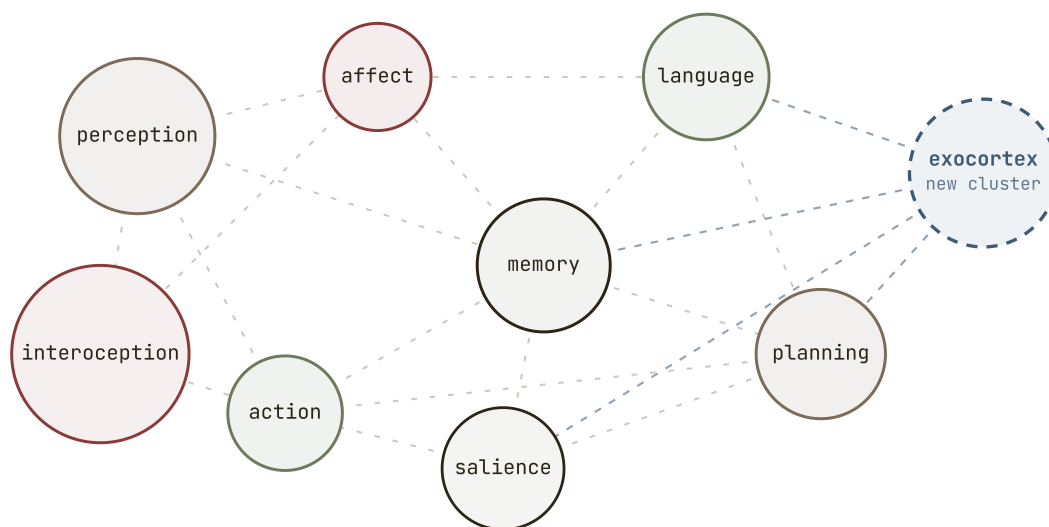
Over longer horizons, neurotechnology may become the only way to meaningfully oversee a superhuman system. Instead of checking every output or line of code, we can couple directly with AI's high-level conceptual map, operating at a matched level of abstraction; and ultimately, we can upgrade our biological neurons to match the capabilities of the machines, creating a **deeper symbiosis** between human and machine. Smartphones have been the primitive version of this symbiosis; BCI and brain augmentation will be a mature evolution of it, through an AI exocortex. In a world with systems a million times more capable than us, high-bandwidth interfaces and brain augmentation are what can keep humans *meaningfully* around, rather than *ceremonially* approving decisions we no longer understand.

If the concept of symbiosis sounds crazy to you, think about it this way: Your brain is not a single, uniform machine, it is *already* a symbiosis of semi-autonomous sub-systems: a brainstem that runs your heartbeat and breathing, subcortical circuits that fire fear and craving a half-second before reason arrives, and a cortex that narrates the whole thing independently. These systems evolved at different times and run on different logic, yet you do not experience yourself as a committee, **you experience a single, unified “I”**. Evolution's way of improving cognition is already to upgrade existing systems and allow new neural networks to integrate into the mind.

Alignment-by-integration would extend that architecture outward. Rather than keeping AI as a separate agent controlled through a narrow interface, BCI could *couple* this external cognitive layer into a new cognitive subsystem of the mind. Alignment-by-control keeps the new intelligence outside the self, a separate thing to be watched and constrained. Alignment-by-integration removes the boundary, the same way your cortex and the rest of your brain share *no boundary you can feel*, a natural extension of Darwinian evolution.

46. One implication is that neural training data should come from carefully selected people whose considered judgments reflect the values we want the model to learn. This may matter less during pretraining, but could be especially important during fine-tuning.

MULTIPLICITY OF THE SELF



phaselock.satpugnet.com

The self already emerges from the coordination of many specialized systems. An exocortex would extend that network outward, adding one more node while the experience remains that of one self.

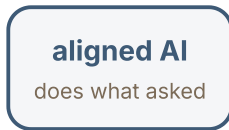
SIDENOTE

The complete picture of human-AI symbiosis requires the ability to write to the brain, on top of reading from it, and eventually, the ability to upgrade our biological circuitry to match the thinking speed of the AI. But we have intentionally barely touched on these topics because they fall outside the core focus of this manifesto, we will nevertheless dive into them further in [Section 20](#) ▶.

In order to be effective, the human-AI symbiosis must happen at a **civilization scale**, not just at an **individual one**. Alignment is a dynamic, pluralistic balancing act, not a static equation to solve. It is never as simple as “align my assistant to me”. True alignment must account for billions of humans, conflicting values, laws, institutions, democratic legitimacy, minority protections, intergenerational interests and non-human conscious beings. While we still do not have a clear plan of how we will handle all those conflicting dimensions, BCI will be a valuable asset in shaping a solution.

A natural question that might arise as we talk about collective alignment is: *who are we actually aligning to?*

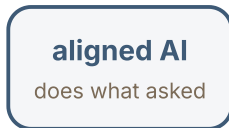
ALIGNED TO WHOM?



To all of us

humanity: plural, contestable

LEGITIMATE



To one of us

one company or one state

STABLE AUTHORITARIANISM

phaserlock.satpugnet.com

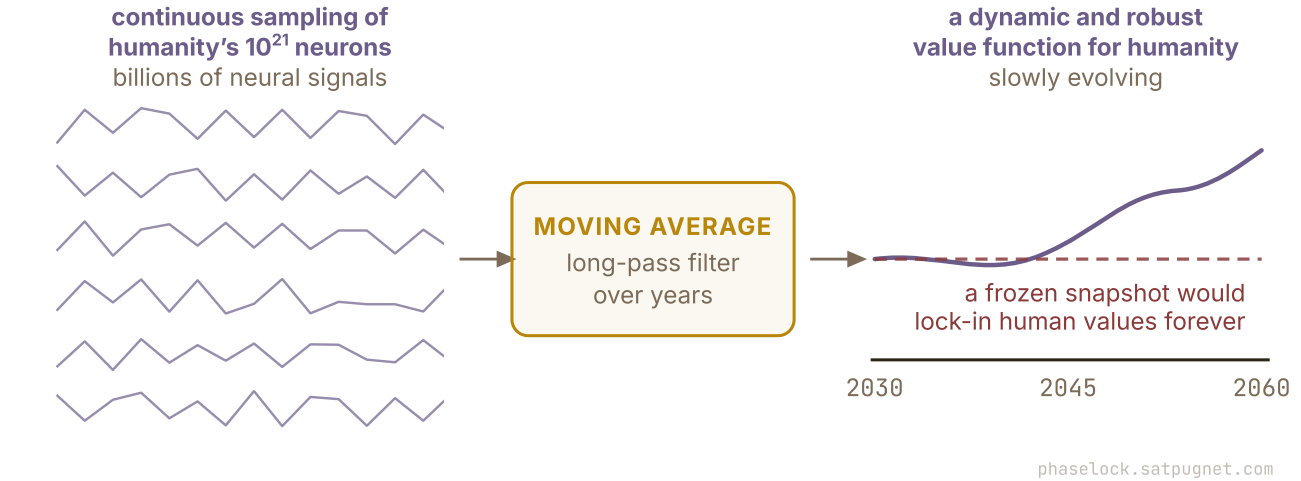
Alignment only guarantees the machine faithfully serves its principal. The very same aligned AI is broadly good if that principal is all of us, and a catastrophe if it is one of us. Successful alignment to the wrong target is its own kind of failure: stable authoritarianism.

In the end, whoever controls the stack will control the target. An AI perfectly aligned via the neural feedback of a small group, a single company, or a nation-state might be problematic in many ways. It would ignore the majority of humans and other conscious beings, giving them no voice into the future. We often treat **accidental misalignment** as the ultimate threat, but **successful alignment to the wrong principal** is just as dangerous.

One speculative practical approach to **defining an objective function at the scale of humanity** would be to *sample the neural activity* of a large and diversified set of humans continuously, *average the semantics* of each individual mind over several years to flatten out short term impulses, and finally use that as a **loss function to steer ASI**. The aim would be to *preserve* the diversity of individual preferences while distinguishing *enduring judgments* from momentary impulses using the highest fidelity source we can access, the mind. Such a signal could provide a **richer empirical foundation** for continuously capturing *humanity's long term goal* (and provide a more robust solution to outer alignment⁴⁷). Crucially, this reference would keep *evolving* alongside humanity over the decades rather than lock future generations into a *fixed* snapshot of our values.

47. Outer alignment is the problem of giving an AI the right objective in the first place, so that what it is trained to optimize matches what we actually want; inner alignment is the separate problem of making sure the trained model truly pursues that objective.

HUMANITY'S VALUE SIGNAL



By sampling every human neurons on the planet and applying a moving average to it, you could theoretically extract a species wide loss function for ASI to use. Such loss function would be much harder to game, it would evolve progressively alongside humanity.

Generally, the premise of most alignment work so far has been to build a separate agent, and then make it obey. How do the slow biological apes constrain the superintelligent digital god? As long as there is an air gap between human and machine, alignment is an **adversarial containment game**.

But high-bandwidth BCI and neurotechnology eventually blur the line between the user and the tool. If your cognitive process is seamlessly integrated with the AI's compute, the AI isn't an "other" that needs to be aligned. It is just an extension of your own cortex, your own self. You don't need to "align" your own arm to your brain; it is part of you. At the limit, BCI transforms the alignment problem from a **containment problem (alignment-by-control)** into an **identity problem (alignment-by-integration)**.

It is important to note though that **higher bandwidth BCI is not automatically alignment-positive**. The same strategy that closes the human-AI bandwidth gap can also be used to create faster manipulation, deeper dependency, and unprecedented surveillance. BCI helps alignment only if paired with strong agency, consent, privacy, and interpretability norms. Like most powerful technologies, it carries **dual-use risks**. We explore those risks further in [Section 22](#) ▶.

Additionally, despite its importance, BCI alone is *necessary but not sufficient*, it is an important lever, but still only one lever. There is still a lot of uncertainty about what exact role BCI will end up playing in this story. But for all the reasons above, it seems highly probable that whichever scenario ends up unfolding, neurotechnology will have a major impact on the world and it is a tool we will absolutely want to have in our arsenal.

I am making some strong claims in this section, and across this manifesto, but we now live in a world where *decades happen in years*. This is *not* just the extended hallucination of a techno-optimist, it is a

thesis grounded in the dizzying new reality we now live in, where technology moves faster than our minds can process. As I write those lines, the **Phase Lock Window** is actively swinging shut, humanity now has little time to build the necessary infrastructure for the New World.

CHECKPOINT III

The first two parts defined the trajectory of BCI. [Part III](#) ▸ examined what happens once it **collides** with AI.

I split it into four sections:

- **Converging Trajectories:** We explored how the walls between these two fields are melting. AI is accelerating our ability to decode the brain, while BCI is opening new sources of data and control for AI. Their trajectories are locking together under the immense gravitational pull of AI.
- **The New Major Modality:** Neural data is about to become a new major training modality, giving models access to cognitive signals upstream of text and behaviour. Neural datasets will unlock billions in AI capital into BCI and accelerate the hardware.
- **The New Interface:** BCI could replace the narrow, sequential channels of keyboards and speech with a continuous stream of intent, context, and feedback, fundamentally changing how humans build and direct AI.
- **Why This Matters for AI Alignment:** Finally, we took a step back and looked at why all of this is critical for our survival. BCI offers several distinct alignment levers: a deeper empirical *map of human values* to inform model design, a *higher-bandwidth channel* that keeps human judgment coupled to AI and ultimately, a path towards human neural augmentation and a deeper *human-AI symbiosis*. Neurotechnology reframes the AI oversight problem from an **adversarial game of containment** towards a **cooperative game of integration**.



PART IV.

The Aftermath

What are the implications for the future?

20 Beyond the Read

Up until now, we have been asking where BCI is going: what role it seems destined to play, why non-invasive reading matters, and why it has a natural synergy with AI. *Now we turn the telescope around*, instead of asking what BCI is becoming, we ask *what kind of future it might create*. First, we will zoom out beyond non-invasive reading and look at the rest of neurotechnology. Then we will move through the geopolitical consequences, the ethical minefield, and the strongest counterarguments against this manifesto. Finally, we will look at what the decade ahead holds before reaching the end of our journey.

Throughout this manifesto, we have been staring at one very specific corner of the map. We focused almost entirely on BCI. Within BCI, almost entirely on non-invasive systems. And within non-invasive systems, almost entirely on reading. In other words, we carved out one narrow slice of the vast neurotechnology landscape and mostly ignored everything else.

This was intentional.

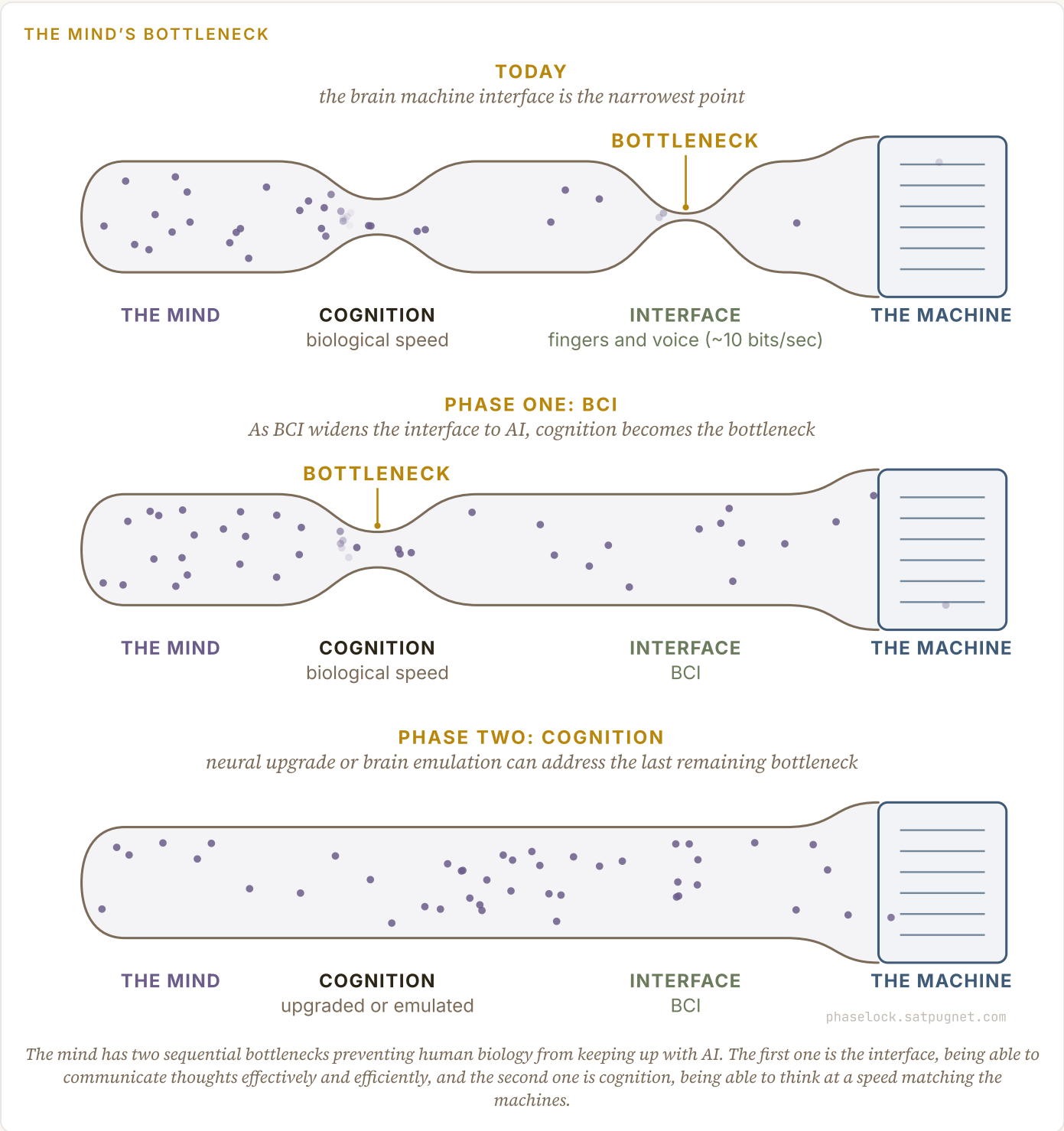
Non-invasive reading is the **commercial spark** that will ignite the entire wildfire. It will act as a massive magnet for capital, talent, and public attention, dragging a reluctant and fragmented ecosystem into existence. It is **upstream** of the rest of the field.

However, the boundaries across this landscape are *incredibly porous*. The foundational pipelines built to decode messy, non-invasive waves transfer directly to clean, invasive data arrays. The hardware teams mastering consumer sensor manufacturing at volume are also building the expertise that can eventually produce surgical implants. The acceleration of non-invasive BCI *will* accelerate the rest of the field.

Every component of this broader neurotechnology matrix will eventually play a vital role in bridging the human-machine gap, maturing on a trailing timeline. While non-invasive reading is starting to hit escape velocity today, I expect consumer-grade **write capabilities** to follow in roughly **five years**, fully **invasive BCI** to hit true scale in **nine years**, **whole brain emulation (WBE)** to arrive within a **15-year** horizon and **neural augmentation** within **20 years**.

Each of those technologies will play a part in addressing the two major bottlenecks of the mind: *how effectively we communicate* and *how quickly we can think*. Over the next few years, BCI promises to substantially improve the communication layer, but ultimately, the biological limits of cognition will remain. Even a perfect brain-AI connection would not let a biological brain keep pace with a superintelligence. Addressing this second bottleneck means **expanding cognition itself**. This is where the broader neurotechnology field comes into play: *neural augmentation* and *whole brain emulation* could extend human cognitive capabilities **beyond their biological limits**, helping close the gap that better communication alone cannot bridge. This manifesto focuses on non-invasive BCI

because it offers the most practical starting point, but neurotechnology’s role in the transition to AGI will reach *beyond* interfaces.



Let’s now explore the blind spots in neurotechnology that we have mostly ignored so far:

Writing to the Brain

For a true human-machine merge, reading the brain is *only half the picture*. The symbiosis we mapped earlier, where an AI seamlessly interacts with you at a higher conceptual abstraction level, implicitly relies on a **two-way street**. In the short term, the machine will just talk back to you through an earpiece or an AR display. But eventually, true symbiosis will require the ability to **write directly back into the cortex**.

Writing is a fundamentally different, and *much harder*, physics problem.

When we read the brain, we exploit a *natural leakage*. Neural activity casts faint electrical, magnetic, and hemodynamic shadows through the skull, which our sensors can catch. Stimulating a specific neuronal population through a closed skull, at the precision required to **cause a thought** rather than disrupt one, is a fundamentally different physical problem than reading. Writing does not have the same leakage in reverse: there is no equivalent path that lets you push a precise pattern into the cortex. Today, our write capabilities are poor. We can broadly stimulate large clusters of neurons, but we are far from elegantly injecting a specific, fine-grained thought. The brain is wildly chaotic, and altering it precisely from the outside remains an immense engineering hurdle.

But the write side of BCI is already making progress. Crude forms of non-invasive brain stimulation are already being used clinically to treat depression, OCD, and other circuit-level conditions. And within a few years, as the massive influx of capital and talent from the reading side of BCI spills over, we will see an **acute acceleration** of the development of brain writing technologies.

Invasive BCI

Beyond the write aspect, we have also spent most of this manifesto focused squarely on non-invasive hardware. Interacting with the brain from outside of the skull. But as we established earlier on [the capability ladder](#) ▶, non-invasive technology has a hard, physical ceiling. The human skull is essentially a massive biological **dampening armor**. Trying to read or write through it makes everything a hundred times harder.

The ultimate, terminal form factor of BCI will absolutely **be invasive**. It will just take longer to get there.

This is the playground of companies like *Neuralink*, *Synchron*, *Paradromics*, *Precision Neuroscience*. By clearing the dura mater and placing electrodes directly on or inside the brain, the physical noise vanishes. The signal in both directions increases by **orders of magnitude** compared to anything that has to pass through bone.

We already know what reading at this fidelity looks like. In a landmark 2021 Stanford study, a paralyzed patient mentally [“handwrote” text](#) at **90 characters per minute** simply by imagining the

motor movements, all captured by an intracortical array. That is the kind of pristine, high-resolution signal you get when you remove the skull from the equation. And that is what will eventually produce the cleanest, highest-fidelity brain interface in the long term.

Invasive hardware will reach consumer scale last because neurosurgery is the ultimate regulatory and psychological bottleneck. But when it finally arrives, it will carry the absolute **highest capability ceiling**.⁴⁸

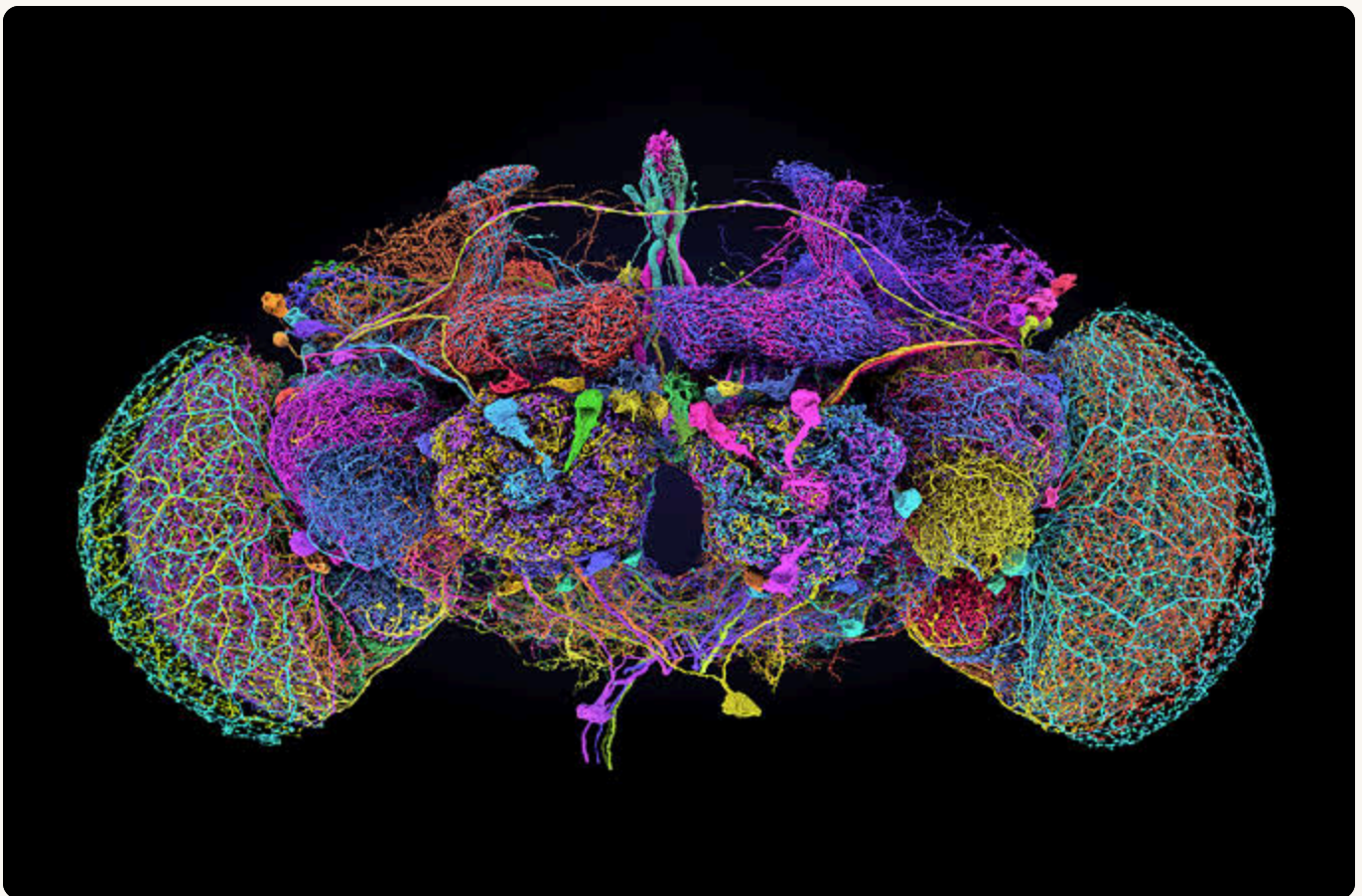


⁴⁸. If you wish to learn more about invasive brain decoding, [this](#) is a good talk on the topic. It breaks down the Stanford study where a BCI decoded handwriting movements into text at record speeds.

Whole Brain Emulation

Taking another step back, there is another area of neurotechnology that we have not discussed yet, but which will certainly play a major role in the future: **Whole Brain Emulation**.

WBE is exactly what it sounds like: the concept of emulating a biological brain directly inside a computer. While this has the distinct ring of science fiction, it is actively transitioning into a concrete engineering discipline. We have already made huge progress in digitally mapping and emulating the entire brains of animals. In late 2024, the FlyWire consortium successfully mapped [the complete connectome of an adult fruit fly](#), detailing the exact wiring of nearly **140,000 neurons** and over **50 million synapses**.



A complete fly brain: In October 2024, researchers published the first complete connectome of an adult fruit fly. Source: [FlyWire](#).

SIDENOTE

In one early demonstration, researchers took a fly's brain data, turned it into an emulated neural circuit, placed that circuit inside a virtual fly body, and [let it drive behavior](#). It moved without a hand-written control system, the virtual mind managed to generate movements directly from the reconstructed brain attempting to run a virtual body.

Whole brain emulation is essentially the other branch of the neurotechnology tree. Much of the capital, talent, regulatory infrastructure, and hardware flowing into BCI also **accelerates WBE** because the two fields share core dependencies: high-resolution neural sensing, models trained on brain data, and standardized methods for mapping the brain.

WBE is also a candidate path to artificial general intelligence. A working human emulation is, by construction, *at least human-level intelligence*, but with the added ability to run faster, operate in parallel, and duplicate at digital scale. Most reasonable forecasts conclude that WBE is unlikely to arrive before native AI systems achieve AGI, but the underlying economics are shifting. A recent thesis evaluating [whether whole brain emulation will matter for the AI transition](#) puts the all-in cost of a human emulation at roughly **\$50 billion** in current scanning and compute terms.⁴⁹

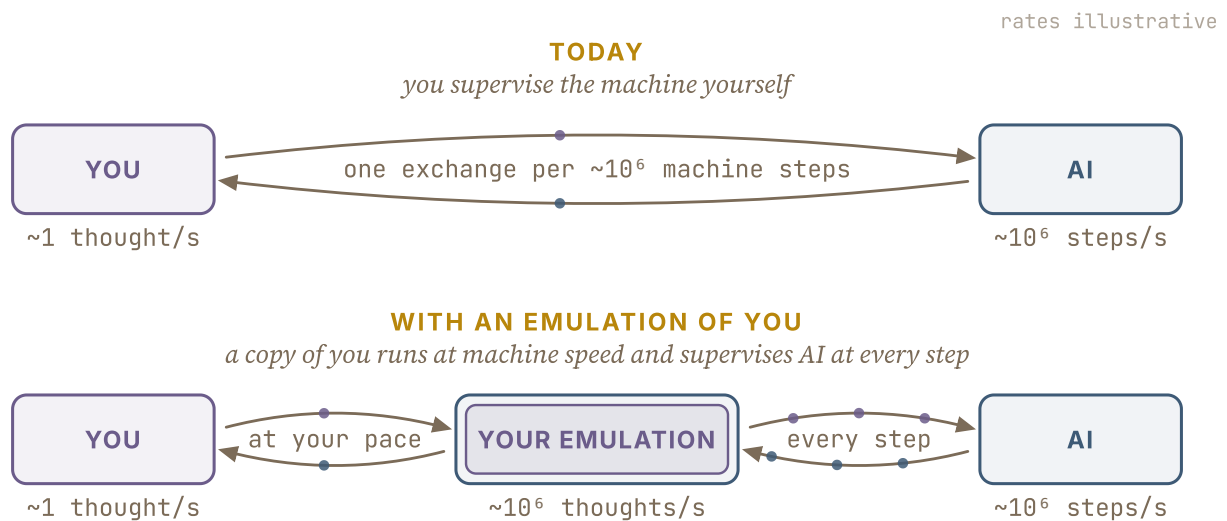
If non-invasive BCI scales exactly the way this manifesto argues it will, the cost of neural data acquisition and processing will plummet. As a result, the **WBE timeline will compress dramatically** right alongside it.⁵⁰

One possible application of WBE for *AI alignment* would be to use it for **supervision at machine speed**. A sufficiently accelerated emulation could *follow* an AI's actions, *evaluate* them against the person's preferences, and *intervene* while its biological counterpart remains outside the immediate interaction. It could then report back at a pace the person can follow. This would offer a potential way to address the *cognitive bottleneck* described earlier.

⁴⁹. For more on WBE's relevance to the AI transition, see [this blog post](#).

⁵⁰. For those wanting to dig deeper into how close we are to turning biological brains into digital substrates, you can explore the roadmap from a [2025 Neurotechnology Workshop](#), which gathered leading researchers.

OVERSIGHT THROUGH ACCELERATED EMULATION



Theoretically, one (or several) accelerated emulation(s) could be used to supervise a superintelligence. Today you oversee the machine at the speed of thought, once for every million steps it takes, an emulation of you could run at machine speed and oversee each step much more effectively.

Neural Augmentation

Looking further into the future, neurotechnology could unlock the possibility of expanding cognition itself. Neural augmentation could *enhance* or eventually *replace* parts of our biological circuitry, **increasing the speed and capacity of thought**. We could make the human mind itself more capable, allowing us to participate *more natively* in a world operating at the speed of machines.

Neural augmentation is still a long way from where we stand today, but given the trajectory of the field, I expect it to become a **reality within two decades**, making mind augmentation one of the *key levers* for long term *alignment-by-integration*.

SIDENOTE

At the limit of our neural sci-fi journey lies **mind uploading**, the possibility of recreating a person's memories, personality, and cognitive processes in a digital substrate. A sufficiently faithful emulation could, in principle, carry aspects of who we are *beyond the limits of biology*. Whether it could be done while *preserving* the original person's "consciousness" or whether it would create a *new* "conscious"⁵¹ individual remains an open question.

But let us come back to the present, where there are still countless challenges to making these visions turn into a reality. Up to this point, we have treated BCI, and neurotechnology, mostly as pure scientific and engineering equations, a matter of physics, compute, and biological ceilings. Yet,

technology *does not scale in a vacuum*, it scales in a messy reality driven by **capital, power, and state interests**. Let us now turn to the geopolitical forces taking shape around BCI.

21 Geopolitical Games

To fully understand how the BCI scenario will unfold, we have to step out of the lab and look at the broader chessboard. As the strategic value of the human-machine merge becomes undeniable, private markets and sovereign states are slowly taking their positions. The race for the ultimate neurotechnology is no longer just a niche scientific endeavor, it is becoming an **economic and geopolitical reality**.

In the past few decades, the dominance of a state, a corporation, or an economic hub has been in large part defined by its mastery over the prevailing technological paradigm. While the coming decades will undoubtedly be defined by artificial intelligence, the interfaces connecting humans to that intelligence will play just as critical a role. Some global powers are starting to recognize this emerging reality and the key role BCI will play over the coming decades.

BCI's increased geopolitical strategic importance will develop over four main phases, with a sharp jump in interest at each step:

- **The Clinical Foundation:** Treating paralysis, stroke, epilepsy, and blindness. This is the bulk of BCI as of 2026.
- **Productivity and Control:** Hands-free computing, spatial interface navigation, silent input, and real-time cognitive workload management. The narrative shifts from medical necessity to human performance and communication with machines. This has already started and will accelerate over the coming years.
- **The AI Data Engine:** Neural data will reclassify from a niche biological dataset into an AI training modality, making the brain a frontier AI asset. This will become viable within a few years.
- **Deep Augmentation:** Closed-loop human-AI collaboration where BCI becomes the new fundamental communication layer with AI. By early next decade, this will be a reality.

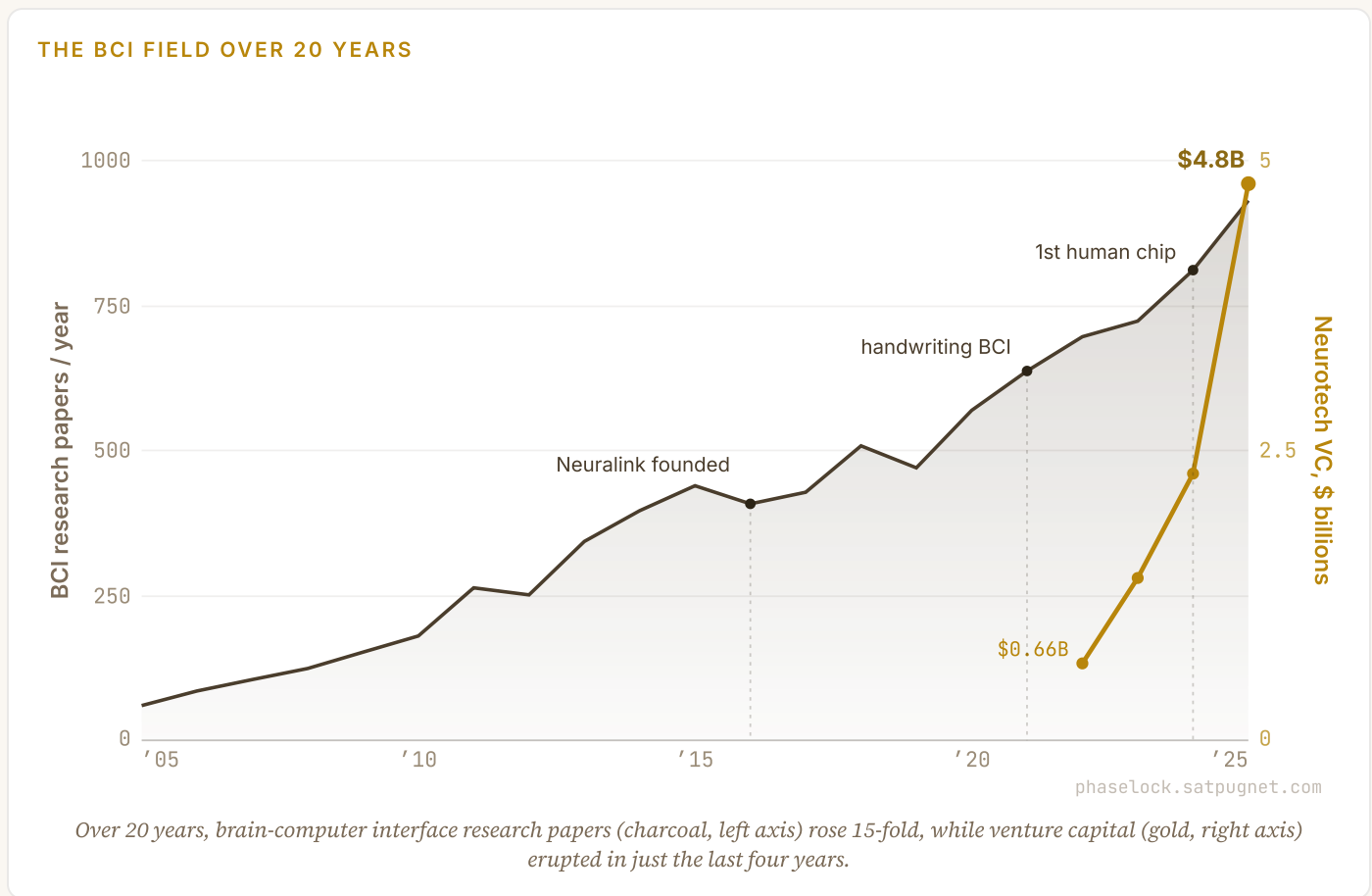
Follow the Money

“Show me the incentive and I’ll show you the outcome.” — Charlie Munger

Money speaks sense in a language all nations understand, and the transition of neurotechnology from academic curiosity to a **high-stakes strategic race** is best captured by looking directly at the money. Private funding into neurotechnology is currently experiencing a rapid consolidation phase.

51. Putting conscious in quotes because I think the term is ill-defined and not very practical.

The macro jump is staggering. Global neurotech venture capital ballooned from \$663 million across 127 deals in 2022 to \$4.8 billion across 140 deals in 2025,⁵² a **7x increase** in just three years.



As this ecosystem consolidates, a dominant tier of well-capitalized corporate players has emerged, holding nearly \$3 billion in combined committed capital:⁵³

- **Neuralink:** Valued near \$9.5 billion following a massive [\\$650 million Series E](#). The company has raised \$1.3 billion cumulatively and rapidly scaled from a single participant to over a dozen human implants across four countries.
- **Science Corporation & BrainCo:** [Science](#) holds \$490 million in total funding after a recent Series C. BrainCo has accumulated roughly \$400 million and filed for a [Hong Kong IPO](#) at a \$1.3 billion valuation.
- **Synchron:** Reached [\\$345 million](#) in total capital, placing its valuation around \$1.0 billion with roughly ten active human implants.

52. Figures from Naveen Rao's [Neurotech Funding Snapshot](#).

53. Sum of disclosed per-company funding, roughly Neuralink ~\$1.29B, Science ~\$490M, BrainCo ~\$400M, Synchron ~\$345M, Merge Labs \$252M, and Blackrock Neurotech ~\$200M.

- **Paradromics:** Around \$120 million raised, advancing its [first human trials](#) through a strategic partnership with Saudi Arabia's NEOM.
- **The Next Wave:** Merge Labs secured a [\\$252 million round led by OpenAI](#) to focus explicitly on semi-invasive BCI. Blackrock Neurotech was stabilized by a [\\$200 million majority stake from Tether](#).

The adjacent neurotechnology stack to BCI has also grown larger in recent years: neuromodulation is already a \$7 billion market, neuroprosthetics sits around \$13.5 billion, and brain implants around \$7 billion.⁵⁴

PARTIAL LANDSCAPE OF BCI COMPANIES

Landscape of Brain-Computer Interface Companies

byFounders

Invasive (by modality)

INVASIVE CORTICAL INTERFACES

Invasive electrodes placed on/within the cortex



MINIMALLY INVASIVE NEURAL INTERFACES

Interfaces that don't require open-brain surgery



NON-CORTICAL NEUROMODULATION

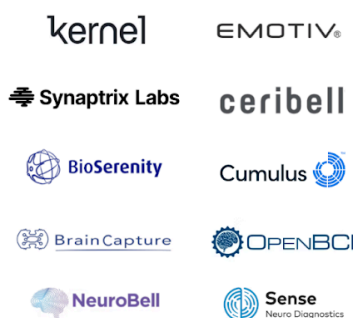
Non-cortical implantable modulation devices



Non-invasive (by function)

NEUROMONITORING

Passively recording and analyzing brain signals



NEUROFEEDBACK

Provides feedback for self-regulation



MODULATION & CLOSED LOOP SYSTEMS

Directly influences brain activity via stimulation



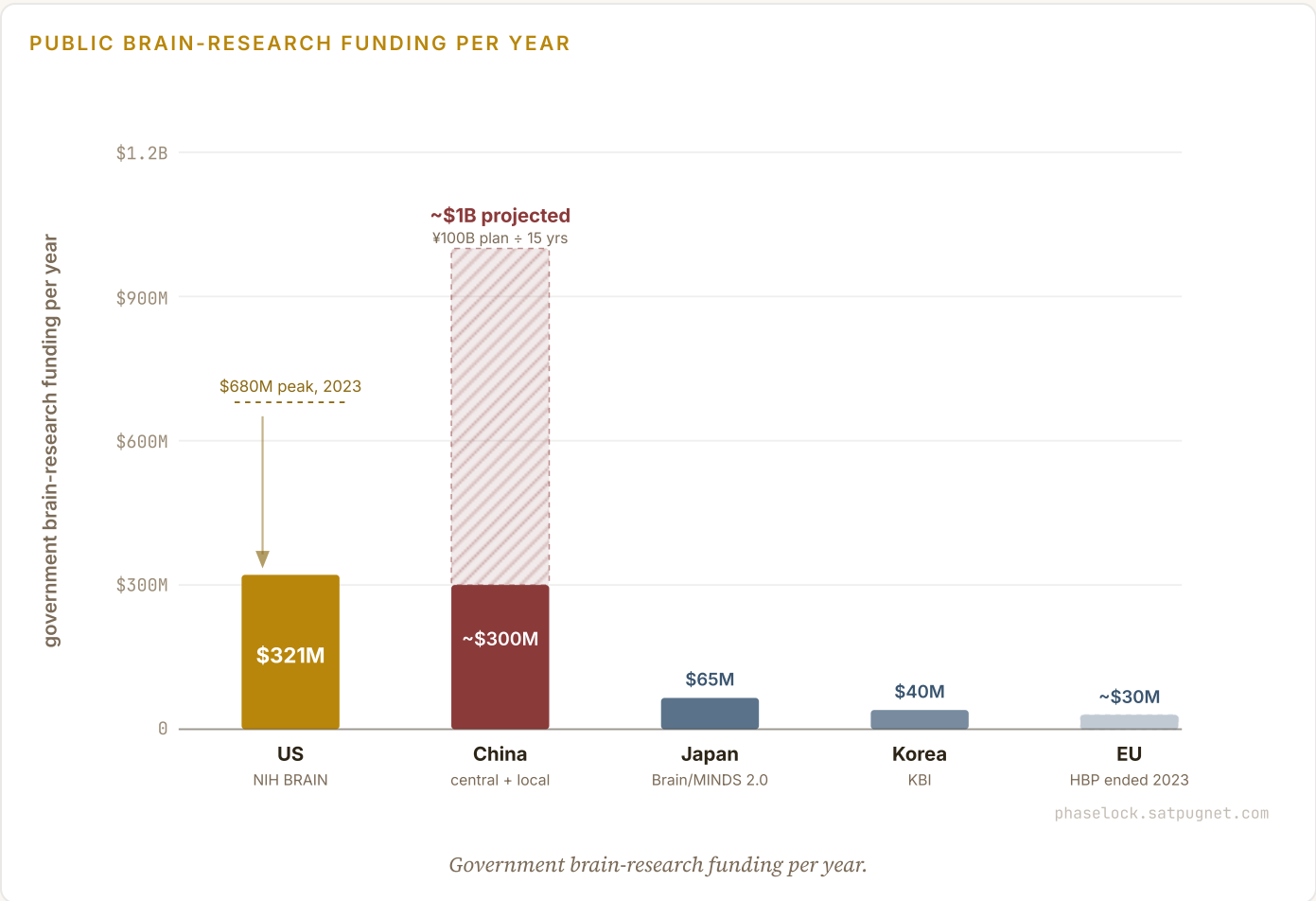
Each year, more companies spawn into existence. Source: byFounders.

When you look past private markets and zoom out to the state level, the public funding balance reveals a *broader* pattern.

54. Sources: [Grand View Research](#) and [MarketsandMarkets](#).

Western public brain-research funding is now over \$4 billion. The US [NIH BRAIN Initiative](#) has deployed roughly \$3.0 billion across more than 1,300 projects. Europe added €607 million through the [Human Brain Project](#) before it ended in 2023. Japan and Korea contributed roughly \$330 million and \$320 million respectively, while the UK is now adding around £200 million across ARIA and UK Biobank brain-imaging efforts.⁵⁵

China is moving at a comparable scale. It plans on deploying a ¥100 billion (roughly \$14 billion) brain-project commitment over 15 years; and has already deployed [\\$750 million from 2021 to 2025](#).



Cognitive Sovereignty

The AI race has so far been a war over compute, fought through chips, fabs, and export controls. BCI opens a **new front**: neural data and access to human cognition. Whoever assembles the largest, highest-fidelity neural corpus will get to train the *best foundation model for the brain*.

⁵⁵. Sources for the Japan, Korea, and UK public figures: Japan's [Brain/MINDS](#), Korea's [Brain Initiative](#), and the UK's [ARIA](#) plus [UK Biobank brain imaging](#).

The familiar dynamics we have seen with AI turning into a strategic technology will follow: concentration, sovereignty, capital, standards, and eventually export controls. But this time, the asset is not text, code, or images, it is the **human mind itself**. Over the coming years, governments will move more aggressively to control and dominate this new frontier.

Cognitive sovereignty will become a major issue: the question of who can collect, model, and act on signals from the nervous system. Neural data is especially sensitive because its meaning can change retroactively, a noisy EEG trace that reveals little in 2026 may expose *critical information* to a better model in 2032.

As a result, who owns neural data, and on whose soil it sits, will become a matter of *national security*. Governments will require it to be stored domestically, as several already do for genomic and health data.

SIDENOTE

The same dynamic already played out with DNA. As sequencing matured, world leaders and their security teams grew wary of leaving biological traces in hostile places, on the theory that a stray sample could later be turned into a targeted bioweapon. This paranoia escalated to the point of cleaning up after presidential visits to avoid the creation of a targeted bioweapon by adversaries (see [Hacking the President's DNA](#)). The threat need not even materialize; the data simply gets more sensitive in hindsight, once a better technology can make use of it.

Neuroethics will also become one of the most consequential policy arenas of the coming decade. The countries that establish the first serious standards for neural data will therefore do more than protect their citizens, they will shape who can build the technology, what it can be trained on, and who ultimately controls it.

The Neural Chessboard

The main actors in the BCI race are currently the United States, China, Europe, and a few other major countries like the UK, Korea and Japan.

As of now, **the US still leads the pack**. A momentum partly due to Silicon Valley and partly to the US government.

On the private side, the US has the strongest startup ecosystem, the deepest venture capital markets, the frontier AI labs, the cloud infrastructure, and companies like Neuralink, Synchron, Paradromics, Precision, and others. On the public side, agencies like [DARPA](#) have funded neurotechnology for decades, especially in the awkward early phase where the science was too advanced for normal grants but too immature for the private market.

However, the US government has been less coordinated than China in recent years.

The US is already developing an **export-control instinct** around neurotechnology. The Congressional Research Service has flagged BCI as a potential control category.⁵⁶ Similar to AI, a new battleground

will open on the BCI supply chain front. It will happen across a vast array of chokepoints: sensors and electrodes, flexible bioelectronics, BCI implants, surgical robotics, biocompatible materials, neural decoding models, neural datasets, regulatory approvals, clinical trial networks, talent, and more. The geopolitical fight will not just be over who builds the best headset or implant. It will be over **who controls the underlying means of production**.

As things stand, the US is still the most likely country to produce the first truly powerful consumer-grade neurotech platforms. However, its fragmented regulatory environment and habit of letting private platforms become de facto governance systems make it equally likely to produce the **first major neurodata scandal**. That is the American model at its best and worst: incredible at creating platforms, terrible at governing them.

China has been especially deliberate recently about BCI, turning it from a scientific ambition into an **operational industrial plan**. In the [14th Five-Year Plan](#), BCI was still mostly buried inside the broader language of brain science, brain-inspired computing, and brain-computer fusion. In the [15th Five-Year Plan](#), it moves up a level: BCI is explicitly named as one of China's strategic future industries, next to quantum technology, biomanufacturing, hydrogen and nuclear fusion, embodied AI, and 6G. The plan is not just to fund labs, but to subsidize the entire manufacturing base to aim for total supply chain dominance by 2030. Today, China already holds 56% of all BCI-specific patents globally, compared to 18% for the US.⁵⁷

The West is losing ground; the US BRAIN Initiative has been contracting its funding while China's Brain Project doubles down.

China is also actively forcing the demand side of the market into existence before perfect consumer products even ship. By establishing [national health insurance billing codes](#) for BCI procedures in 2025, they laid the groundwork for the world's first commercially approved implantable BCI ([Neuracle NEO](#)) to hit the market in early 2026, beating Neuralink to the regulatory milestone.

China has shown the same pattern before in electric vehicles, batteries, drones, solar, and advanced manufacturing: once Beijing decides a technology is strategic, it can align capital, factories, standards, suppliers, and customers with unusual speed. The US still leads the BCI race today, especially through its frontier startups and AI ecosystem. But China is now moving from *catching up* to **mobilizing the whole country**.

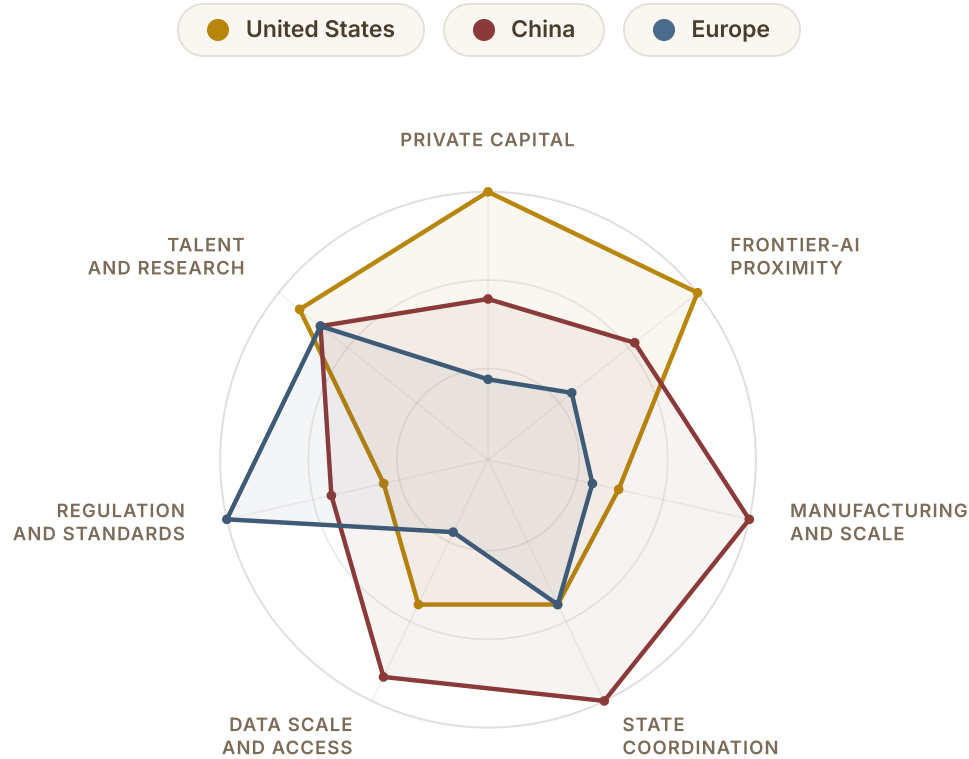
Europe, by contrast, is playing the role of **the regulatory superpower**. The EU possesses strong neuroscience institutions, robust public ethics, and deep institutional trust, but a chronic inability to build dominant consumer tech platforms. Through comprehensive frameworks like the AI Act and global pushes for constitutional neurorights, Europe is actively setting the global ethical baseline for mental privacy, cognitive liberty, and freedom of thought.⁵⁸ This gives Europe tremendous

56. Commerce Department's Bureau of Industry and Security [requested public comment on export controls for BCI](#) as an emerging technology.

57. Source: [JMIR Rehabilitation and Assistive Technologies](#).

normative legitimacy, but not necessarily hard power. Europe is positioning itself to **write the rules for neurotechnology**, while inevitably importing the American and Chinese platforms those rules govern.

SITUATIONAL AWARENESS



phaselock.satpugnet.com

The US stretches toward private capital and frontier-AI proximity; China toward manufacturing, state coordination, and the scale of neural data it can collect; Europe toward regulation and standards.

Defense and Dual-Use Matter

BCI is a dual-use technology.

And the US and Chinese militaries are well aware of this; they have been subsidizing the field for years in the hope that it would become mature enough to be used in warfare.

The same headset that lets a designer move faster through a spatial interface could let a drone operator control a swarm without touching a joystick. The same neural decoder that helps a paralyzed patient speak could let a soldier communicate silently in a warzone.

58. The EU AI Act's [Article 5](#) bans manipulative and subliminal systems, the closest existing regulation for mental privacy. Although, it is far from enough.

This is not just speculation; defense agencies have been funding neurotechnology for decades, with particular interest in human-machine teaming, soldier monitoring, rehabilitation, training, and command latency. In war, shaving seconds off the observe-orient-decide-act loop⁵⁹ matters a lot.

There is already serious policy attention here. A 2025 [GPPI report](#) argues that military actors are deeply involved in US and Chinese neurotechnology ecosystems.⁶⁰ The [ICRC](#) also flagged military BCI as a new international humanitarian law issue in 2025.

One of the clearest tells is **the cap tables**. Synchron's [2025 round](#) brought together traditional tech investors with the Qatar Investment Authority, Australia's National Reconstruction Fund, and In-Q-Tel, the CIA's venture arm. Spy agencies and sovereign wealth funds investing so intently in frontier technology usually tells a clear story which goes beyond the medical device.

Over the next few years, neurotechnology will move beyond healthcare and consumer electronics to turn into **cognitive infrastructure**. And just as with AI and semiconductors, any critical infrastructure technology eventually becomes a **national security priority**.

This is where the ethical questions become *impossible* to avoid. As we start altering perception, intention, attention, memory, agency, and choice; we are touching **human identity at its core**; and when the fundamental substrate of a technology is the human mind itself, one can easily conjure a dozen *dystopian* scenarios for the future. So what if things go wrong?

⁵⁹. The observe-orient-decide-act (OODA) is a rapid decision-making framework developed by military strategist John Boyd; the side that completes it faster seizes the initiative. An interface at the speed of thought can compress the delay significantly.

⁶⁰. See "[Action Potentials: Neurotechnology, Brain-Computer Interfaces and the Future of Defense](#)".

22 Mind Reading at Scale: What Could Possibly Go Wrong

“Technology is neither good nor bad; nor is it neutral.” — Melvin Kranzberg



The fall of Icarus: Icarus flew too close to the sun and the wax melted, the wings worked but he aimed them wrong. A technology can work beautifully and still be pointed in the wrong direction.

For as long as humans have existed, the skull has been the only truly private room in the universe. This may soon be a concept of the *past*.

Every technology eventually crosses a line where the question stops being “*Can we build it?*” and becomes “*What happens if it works?*”

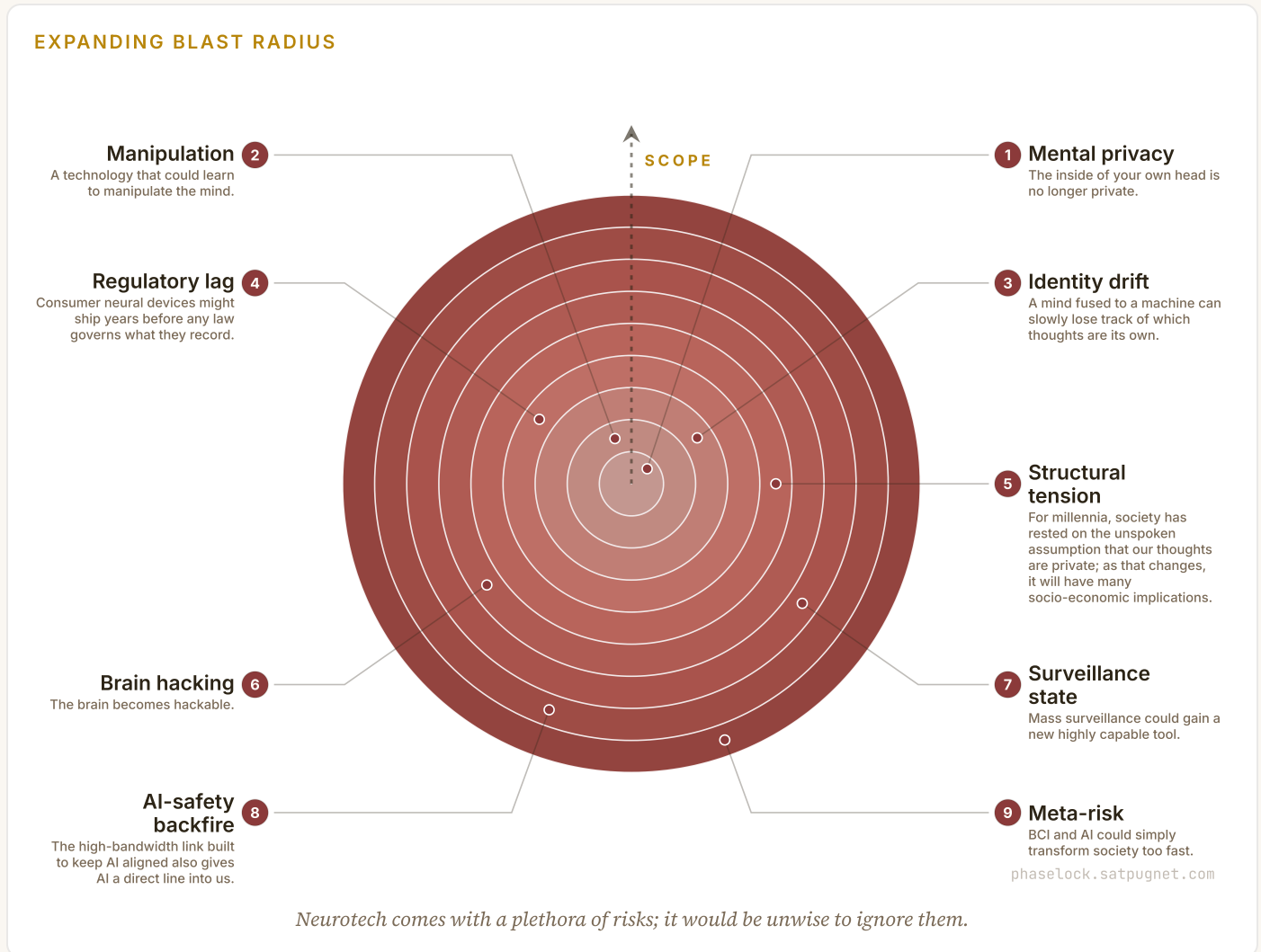
BCI is approaching that line.

And with neurotechnology, it does not take much imagination to see how it could go wrong. BCI is not another app, another platform, or another productivity tool; it reaches into the most intimate layer of human life: **the mind itself**.

That is exactly what makes both the upside and the risks so *enormous*.

Like AI risk, the risks of neurotechnology are poorly understood, and for most people the best mental models they have come from science fiction. Given the breadth of the topic, a serious treatment would be a manifesto of its own,⁶¹ we will therefore only be scratching the surface here.

The Neurotech Risk Surface

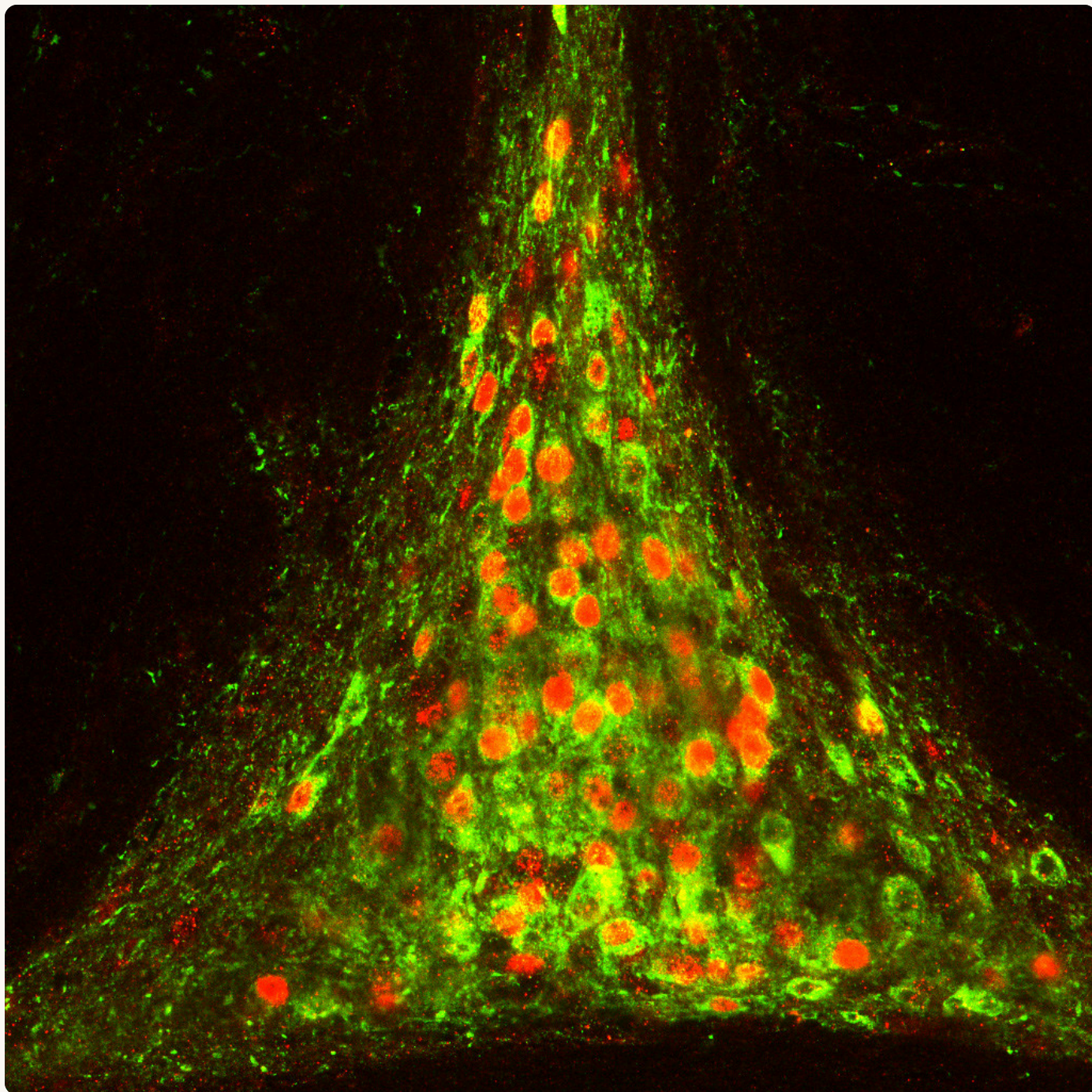


The risks below are ordered by expanding blast radius; from what neural interfaces can reveal and manipulate within an individual mind, to their effects on identity, society, and ultimately the trajectory of AI:

Mental privacy. Neural data is a leaky signal: it can reveal more than the user intends. As decoders improve, a neural signal that today tells you whether someone is thirsty can later be used to infer

61. Further reading: [Yuste et al.](#), [Ienca and Andorno](#), and Nita Farahany's [The Battle for Your Brain](#).

depression risk, political leanings, attraction, addiction vulnerability, or job performance. Harmless data today can become intimate data tomorrow, and once collected, it is hard to take back.



Thirst neurons in a mouse brain. This is what thirst looks like before it becomes a sentence. Source: [NIH](#).

Manipulation. Once BCIs mature, they will stop being passive sensors and writing will become possible. They will be able to recommend, nudge, stimulate, suppress, or amplify thoughts and emotions. The main failure mode becomes not only malicious actors hacking your thoughts, but also

your attention, mood, and choices being gradually mediated by a system optimizing for something you did not choose, with access to your raw cognitive machinery. We have run a version of this experiment before, with recommendation algorithms on social media restructuring public discourse while holding nothing but your click history. Give those same algorithms the error signal straight from your limbic system, and the result could be *dangerouslysuperhuman persuasion*.

Identity drift. A BCI that augments someone can also create dependency, dissociation, and over-trust in machine output. With a closed-loop AI operating this close to the mind, *identity itself blurs*. Did I choose that, or did the model predict and complete me? Did the stimulation improve my mood, or alter my preferences? Am I becoming more myself, or more optimized? These questions have no clean answers, and the drift they describe will be hard to control.

Regulatory lag. Regulation is nowhere near ready, and the ceiling problem makes it worse: when you do not know what a technology will eventually decode, the rules have to cover the extreme futures, not just the current demos. If a few minutes of EEG could one day reveal whether someone committed tax fraud, is it acceptable for that data to sit in a training corpus today? Consent has a time-machine problem: you find out what you gave away only after the models improve. Genomics already screened this movie, 23andMe's bankruptcy turned the DNA of 15 million people into an asset for sale.⁶² Consumer neurotech is currently worse, in an audit of 30 companies, 29 granted themselves effectively unrestricted access to neural data, nearly all reserved the right to transfer it, only one in five even mentions encryption, and exactly one rules out sharing it altogether.⁶³ Meanwhile the deepest questions have no case law at all. For instance: if a BCI logs your P300⁶⁴ recognition response to a crime scene photo, can it be subpoenaed? Does the Fifth Amendment cover a brainwave?⁶⁵

62. When 23andMe went bankrupt in March 2025, the DNA of its ~15 million customers became an asset for sale, going [for \\$305 million](#).

63. Source: [Safeguarding Brain Data](#).

64. The P300 is an involuntary brainwave that peaks about 300 milliseconds after a stimulus and spikes when a person recognizes something familiar. It is the basis of concealed-information tests: the brain flags recognition before the person decides to reveal anything.

65. The Fifth Amendment protects you from being forced to testify against yourself. Whether a decoded brain signal counts as testimony, or as physical evidence like a fingerprint, is a question no court has faced.

ONE RECORDING, RE-READ OVER TIME

2026. You wore a **\$99 headband** to sleep and focus better, and agreed to one thing: let it **score your focus**.

RE-READ BY 2026

Focus and fatigue

locked in, or fading ·

18% confidence

Stress and mood

out of reach

Your real reactions

out of reach

What persuades you

out of reach

Politics and values

out of reach

Mental health and addiction

out of reach

Attraction and desire

out of reach

What you will do next

out of reach

So far it tells your headband what you agreed to: **a focus score**. Nothing you would not have shared.

RE-READ BY 2030

Focus and fatigue

locked in, or fading ·

82% confidence

Stress and mood

what winds you up, what calms you ·

68% confidence

Your real reactions

the reaction behind the polite face ·

50% confidence

What persuades you

what moves you, and how to sway

you · 31% confidence

Politics and values

who you are most likely to vote for ·

10% confidence

Mental health and addiction

out of reach

Attraction and desire

out of reach

What you will do next

out of reach

Same recording, better models. It now reads **politics and values**, which you never handed over.

RE-READ BY 2035

Focus and fatigue

locked in, or fading ·

91% confidence

Stress and mood

what winds you up, what calms you ·

81% confidence

Your real reactions

the reaction behind the polite face ·

68% confidence

What persuades you

what moves you, and how to sway

you · 58% confidence

Politics and values

who you are most likely to vote for ·

47% confidence

Mental health and addiction

depression, anxiety, dependency,

early · 39% confidence

Attraction and desire

who you are drawn to ·

26% confidence

What you will do next

your next move, before you make it ·

6% confidence

The recording you gave for a focus score now infers **what you will do next**. You never consented to that.

phaseLock.satpugnet.com

You consented to a focus score but the recording can confess a lot more.

Structural tension in society. Human society has run for millennia on the assumption that thoughts are private. Courts, marriages, negotiations, elections: all of them depend on the fact that no one can check what you really think. AI already challenged our monopoly on intelligence, sending ripples and *fractures through society*. BCI is about to challenge the privacy of the mind. If that shift arrives faster than society can rebuild itself around the new reality, the damage could be severe. There is also the risk of cognitive inequality and concentration of power: whoever controls the brain-and-AI stack

controls the interface to intelligence itself, which could widen the gap between social classes, age groups, and countries.

Brain hacking. As strange as it sounds, everything connected can be attacked, and a read-and-write brain device would be the most intimate attack surface ever shipped. The threat model is ordinary cybersecurity with new stakes: signal interception, spoofed inputs, malicious neuromodulation, denial of service. A ransomware would feel differently if what it encrypts were your ability to think clearly. Every security failure of the last 30 years can be replayed here, against a piece of hardware you cannot uninstall.

The surveillance state. Every surveillance technology has followed the same trajectory: *smaller, cheaper, easier to hide*, every year. Cameras and microphones did it; neural sensors will too. In the wrong hands that means emotional surveillance, dissident profiling, interrogation, and eventually thought decoded at population scale. The polygraph⁶⁶ was junk science and still got a century of use; now imagine a version that actually works, deployed at scale by an authoritarian state.

The AI-safety backfire. On the AI side, the sharpest risk is that neurotechnology does not help with alignment at all and only augments AI capabilities. Humanity's inner life becomes one more pretraining corpus while the technology grows ever harder to control. And even if BCI does help steer and align AI, there is a risk that we align it on a subset of minds not representative of humanity. Lastly, if neural data becomes as valuable as we expect, *predatory behaviors* may emerge to collect it at scale.

The meta-risk. Underneath all of these sits something larger: the risk that humanity is simply not ready for this shift of its reality. We may end up governing twenty-first-century access to the mind with twentieth-century categories, because those are the only ones we have. AI was the first blow to human identity; BCI is the second, and it raises the question directly. What is a human being when the boundaries of the mind are negotiable? What is the self? What is our role in the world that follows? None of these questions will be easy to answer.

Neurotechnology will force some of the most profound questions humanity has ever had to answer, on the timeline of a product cycle.

How This Goes Right

So if the risk surface is this large, **why build BCI so fast?**

⁶⁶ The polygraph, or “lie detector”, measures stress as a proxy for lying. Its scientific basis is [shaky at best](#), yet it was used on millions anyway.

The honest answer is that, were AI capabilities not growing exponentially, I would advocate for slowing down the development of neurotechnology to give us more time to get it right. But as the world stands, building BCI has now become a **necessity**. The age of machines is upon us, and BCI is one of the few technologies with a real chance of making that transition survivable for humanity, one of the few chances we have to do alignment-by-integration. At this point it is not just an aspirational techno-utopian dream, it is becoming **existential**.

Building it right will be extraordinarily hard, as every scenario above should have made clear. But it is doable, good outcomes are possible; they simply have to be treated as an *active design problem* rather than the *default outcome*. And because we still only half understand the technology, the frameworks we build around it have to evolve dynamically.

We should start with **architecture**, because the strongest guardrails are not legal but technical. Raw neural data should leave the device as rarely as possible: process locally, transmit intent rather than brainwaves, export features rather than raw traces, train through federated learning and differential privacy so models improve without the signal ever pooling on anyone's server. Where large corpora are genuinely needed, they can live in independent data trusts that audit and license access, rather than letting any single platform accumulate humanity's inner life. A law can be repealed by the next administration; but an architecture has to be re-engineered. **Privacy enforced by physics is the only kind that survives a bad decade.**

We have already worked *miracles* with the internet and with biotechnology, we can pull it off once more.

The same logic applies to **business models**. Social media did not become an outrage machine because its engineers were malicious; it became one because advertising was the revenue model and attention was the product. The cheapest way to avoid replaying that mistake at the neural layer is a bright line: brain data and advertising must never meet.

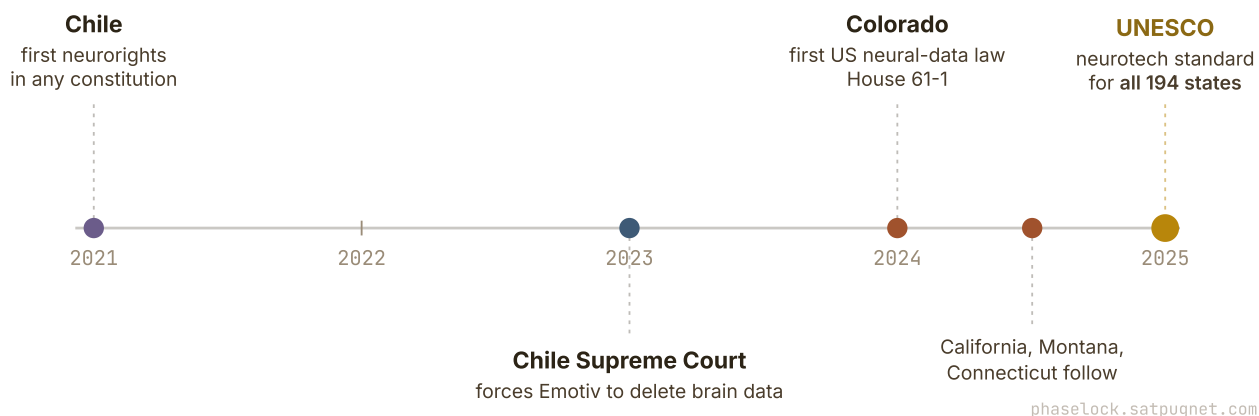
Consent has to grow up too. A checkbox at install cannot govern a data stream whose meaning changes as models improve. Neural consent needs to be granular, auditable, revocable, and, above all, perishable: it should expire unless renewed, because next year's models will read what this year's cannot. That is the answer to the time-machine problem. If you cannot take back what you streamed, then streaming has to become a running conversation rather than a one-time signature.

The **law**, for its part, needs one conceptual upgrade more than it needs a thousand new pages: neural data is not app data, not wearable data, not even ordinary biometric data. It deserves its own category at the very top of the sensitivity hierarchy, with strict purpose limitation, no resale, no secondary inference (the focus tracker that covertly scores your politics should be illegal, not just creepy), no employer or insurer access by default, robust deletion rights, and hard limits on what can be trained from it. Above all, *cognitive liberty*, the right to mental self-determination, must be recognized as a fundamental right rather than a nice idea in an ethics paper.

SIDENOTE

And here, unusually, there is good news. Chile [wrote neurorights into its constitution](#) in 2021. Colorado, California, Montana, and Connecticut have [extended their data-privacy laws to neural data](#), and their definitions are becoming the de facto American standard. UNESCO's [Recommendation on the Ethics of Neurotechnology](#), adopted by its member states in November 2025, is the first global normative instrument for the field: soft law, but a shared vocabulary. The EU AI Act's guidance already cites brain-computer interfaces by name in its rules on [prohibited manipulation](#). In roughly two years, neural data has gone from no specific protection anywhere to binding or near-binding coverage for some 50 million people. Regulation usually arrives years after the scandal, for once, it has a chance of arriving before.

THE NEURAL-RIGHTS TIMELINE



Evolution of the neuroright landscape.

Two lines deserve to be drawn thicker than all the others. First, the write channel is sacred. **Reading the brain is dangerous; writing to it is civilization-level dangerous**, and the standards should treat it accordingly: the closer a system comes to changing a mental state, the higher its burden of proof should be. Second, **the brain must be treated as critical infrastructure**: threat-modeled, red-teamed, secure by default, and protected internationally by red lines drawn before escalation rather than after. No neural interrogation. No compulsory brain monitoring. No covert neuromodulation. And no neuroweapons. Humanity eventually built such norms for biological and chemical weapons; this time we could build them before the first use instead of after it.

There is also a failure mode on the other side though, if we regulate this field into the ground, we will *not* get more safety; we will freeze BCI precisely during the decades when it matters most for navigating the transition to advanced AI. And in addition this hands the frontier to whoever does not ask permission. The goal is not to be slow, it is to be deliberate about how we build the technology. We have to move fast but intentionally: humanity's future is at stake, we cannot afford to pause here, and we cannot afford to get it wrong.

Alongside all of this, we need to raise **public literacy**. Just as AI did, neurotechnology is about to enter the global conversation, and it should be something people actually understand. A society with

an accurate mental model of what these devices can and cannot do is better equipped to choose well, it also becomes harder to panic, and harder to lull.

Kranzberg's law⁶⁷ is usually quoted as a warning. This technology will not be good on its own, and it does not have to be bad. It will be exactly as good as the standards we set while it is still young enough to accept them, and that work is needed right now.

Ultimately, if neurotech treats neural data the way ad-tech treats browsing data, *it deserves to fail*. Every guardrail in this section, the architecture, the consent, the law, the red lines, should exist to keep the technology on the right side of history.

And beneath all the policy sits a choice no regulator can make for us: **deciding what we want to preserve about being human**. Is it *agency*? *Authenticity*? *Intelligence*? *Mental self-determination*? The *privacy* of an inner life? We should name those things deliberately and design toward them, or they will be defined for us.

BCI has the potential to help humanity keep its hands on the wheel as AI grows ever more powerful, while unlocking countless other capabilities along the way, all through a technology we could barely dream of a decade ago. But it is also one of the most *dangerous* technologies humanity has ever brought within reach, and it should be treated accordingly.

Let us now turn to what the coming decade could look like if we make the right choices.

So far, I have worn the hat of the **cautiously optimistic futurist**, mapping the most plausible path forward, but any serious thesis must survive its *strongest objections*. So let us swap hats: if this vision hits a wall, where will it happen, and what could derail it?

23 The Case Against the Case

“I am not entitled to have an opinion unless I can state the arguments against my position better than the people who are in the opposition.” — Charlie Munger

There is no shortage of reasonable objections to this thesis, ranging from the physics of the human skull to the information theory of foundation models and the basic economics of hardware scaling. To earn the right to be taken seriously, I cannot just swat at strawmen; I have to face the strongest objections head-on.

Here are the 13 objections that should be addressed.

01. Non-invasive may have a fundamental signal ceiling that scaling cannot break.

⁶⁷. Melvin Kranzberg's first law of technology, and the epigraph of this section: “Technology is neither good nor bad; nor is it neutral.”

The skull blurs and attenuates neural activity, leaving a noisy, underdetermined signal. Scaling can extract hidden structure, but it cannot recover information that physics has erased.

RESPONSE We do not know where that ceiling sits. This is ultimately an **empirical question**, and the field has not yet scaled hardware, data, and models far enough to answer it.

02. Non-invasive research may divert resources from the higher-potential invasive path.

Capital, talent, and institutional attention are finite. If invasive systems have a much higher capability ceiling, directing the field toward easier-to-market headsets could delay the path with the greatest long-term potential.

RESPONSE That argument made more sense in 2015 than it does in 2026. Hardware advances, data standards, decoder architectures, and clinical infrastructure increasingly transfer across modalities. Capital and talent drawn into non-invasive BCI *strengthen* the broader field, even if invasive ultimately wins. More importantly, for the foreseeable future, non-invasive is the only path with a **realistic chance** of escaping the lab and reaching millions of people.

03. The brain may not be the same kind of scaling problem as the internet.

Brain data is noisier, more individual, and structured across many time scales; its labels are also far weaker. The Bitter Lesson may not transfer cleanly.

RESPONSE As with the signal ceiling, this is a question *only scale* can settle. So far, the evidence points in the right direction: ZuCo, DEAP, AllJoined, [BrainLM](#) ▶, and Banville et al. all show performance improving with more neural data across multiple modalities. [ENIGMA](#) ▶ reached state-of-the-art EEG image decoding on both consumer and research-grade hardware, using under 1% of the parameters of prior systems and only 15 minutes of calibration for a new subject. There is **early evidence**, not just aspirational hopes.

04. BCI may lack the universal self-supervised⁶⁸ objective that made AI scaling work.

Language modeling scaled because predict-the-next-token is stupidly universal: every paragraph on the internet is a labeled training example for free. Brain decoding has no such clean objective; most current approaches train task-specific decoders that do not transfer.

RESPONSE The shape of the emerging objective is starting to look like **multimodal world-brain prediction**. Given a neural signal plus context (eye movement, audio, behavior, task state, screen content, body pose), learn the latent representation that predicts the missing channel. Less elegant than next-token prediction, but *not obviously* less scalable; indeed, AI has since scaled successfully on

⁶⁸ Self-supervised learning (SSL) trains models using targets generated by the data itself rather than human labels. The canonical example is predict-the-next-token in language modeling; every paragraph on the internet becomes a labeled example for free. SSL is the reason modern AI scaled so cleanly.

messier objectives, from masked-image modeling to contrastive learning across images and text. And the unit of training data, “brain signal plus context”, is exactly what consumer hardware will produce by default.

05. Subject variability may kill foundation-model economics.

Brains differ in anatomy, plasticity, and biochemistry. A decoder trained on one person rarely transfers cleanly to another.

RESPONSE This is the core bet [Section 10](#) traces in detail: that pretraining across many brains *collapses the marginal cost* of decoding a new one. If pretraining works for brains the way it works for language and vision, personalization becomes a short calibration step rather than a bespoke training process.

06. BCI has no parallel consumer market subsidizing its manufacturing iteration.

Modern AI scaling was carried for over a decade by adjacent multi-billion-dollar markets that drove the underlying silicon, such as video games and crypto. They paid for the Nvidia GPU R&D that AI eventually needed.

RESPONSE The world is a densely connected graph. As a result, much of the hardware BCI will require is already being subsidized by adjacent industries: smartphone sensors, consumer-health wearables, mature CMOS, cheap optics, AR/VR, and decades of declining sensor costs. This is also going to be greatly amplified by the incoming rise of robotics. Some genuinely new technology will still be required, but most of the stack can be assembled from increasingly capable **off-the-shelf hardware**.

07. The consumer-EEG graveyard.

Muse, Emotiv, NextMind, CTRL-labs, a handful of others: the last consumer non-invasive wave produced demos, acquisitions, and shutdowns, not a sustained product category.

RESPONSE The earlier wave tried to build products before the underlying stack was ready. There were no foundation models for the brain, no multi-modal sensor stacks, no shared protocols, no AI labs hungry for data, no cheap consumer-grade sensors. Just like most technologies, **it does not work until it does**; it is all about timing. The case today is not simply that “this time is different”, but that several missing pieces now *exist*.

08. Neural data may not be valuable enough to frontier AI labs.

Labs already have enormous text and video corpora and are increasingly relying on synthetic data, reinforcement learning, and simulated environments. Noisy, expensive, consent-sensitive neural data may add too little marginal information to justify significant investment.

RESPONSE Neural data is potentially valuable because it **captures aspects of cognition** that text, video, and synthetic data can only expose *indirectly*: raw information processing from the brain. That could improve both capability and alignment ([Section 17](#) ▶, [Section 19](#) ▶). And as frontier labs approach the data wall and alignment becomes increasingly central, genuinely new modalities with a chance to address those issues become more valuable. But this remains an empirical bet: labs will need to test whether neural data produces measurable gains on capability and alignment.

09. Neural data may be too private, too regulated, or too socially radioactive to become a frontier training modality.

People accept watches that log sleep and heart rate. They may feel differently about a device tracking attention, recognition, or emotional state. One sufficiently ugly scandal could set the field back years.

RESPONSE Exactly why on-device preprocessing, differential privacy, and institutional legitimacy have to be designed into the technology from the start. The wrong version of this future would be extractive, opaque, and dangerous; the right one would be *opt-in*, *auditable*, and clearly *valuable* to the person generating the data. Once again, if the field treats neural data like ad-tech data, it deserves to fail.

10. The real human bottleneck may be comprehension, not communication bandwidth.

A faster pipe does not make the biological brain faster at understanding, judging, or verifying superhuman reasoning. A thousand-fold increase in input bandwidth could simply become a firehose into a bathtub. If comprehension is the binding constraint, BCI tweaks the wrong knob.

RESPONSE In the near term, BCIs will increase the practical bandwidth between humans and machines by: reducing the meaning lost in translation, enabling AI to adapt explanations to a user's understanding in real time, and operating at the right level of abstraction rather than dumping raw reasoning into the mind. However, fundamentally, communication bandwidth will not increase human thinking and processing speed; and as AI's thinking speed accelerates, BCI alone will not allow the biological brain to keep pace with AI. This manifesto's claim is that BCI will only be the first step towards closing this gap, it will catalyze advances across neurotechnology that will then collectively help close the human-machine gap. As [the mind's bottleneck](#) ▶ illustrates, communication bandwidth is only one of two bottlenecks, and it will be the first and most tractable one to address. Nevertheless, over a longer time horizon, neurotechnology will also offer a path to addressing the second bottleneck: cognitive speed; through neural augmentation and brain emulation, we will have a path to overcome the biological limits of cognition itself.

11. A better human-machine interface does not solve alignment on its own.

At best, it improves how clearly humans can express intent, preference, and correction. But alignment is ultimately about what the system is optimizing for, and a more capable interface may be

only tangential to that problem. A misaligned system does not become safe simply because it can understand us better.

RESPONSE A better interface is not the complete alignment solution, but it is nonetheless a significant part of it. Neurotechnology could improve reinforcement learning, provide richer signals of human intent and preference, surface confusion or disagreement, and overall **help train models** that better understand *conflicting* and *complex* human values (see [Section 19](#) ▶). BCI, and neurotechnology more generally, is one of the only technologies with a real shot at enabling *alignment-by-integration* before the Phase Lock Window closes, its development should therefore be prioritized accordingly.

12. A neural interface may strengthen the machine more than the human.

The AI could use access to attention, emotion, uncertainty, and preference to model us with extraordinary precision, while our ability to understand or supervise it improves only marginally. The same channel meant to preserve human control could instead become a powerful tool for persuasion and manipulation.

RESPONSE This is a **real danger**. Better access to the human mind can improve *oversight*, but it can also make AI *persuasion* far more potent. The solution is to not assume neurotechnology will naturally favor the human; but instead to be **intentional** in our development and design of the technology to *mitigate* those risks.

13. Neural feedback may optimize for impulses rather than values.

Fast neural reactions are not the same as reflective preferences. They contain novelty, fear, arousal, bias, addiction, and immediate reward. Treating them as biological ground truth could turn RLHF's existing failure modes into something much stronger: models optimized not for what humans endorse, but for what most effectively lights up the human reward system.

RESPONSE Neural signals should indeed *not* be treated as unquestionable ground truth for training, their value is as one signal among many. Used carefully, they may help detect confusion, error, or hidden disagreement; used naively as a reward target, they become another Goodhart's failure⁶⁹, optimizing for stimulation rather than human values. The solution here is to remain mindful of the way we make AI ingest this new signal.

We are currently navigating **treacherous waters**. Blindly following a roadmap drafted with information from 2026 is a recipe for disaster. The map is not the territory, and any attempt to project decades of technological evolution is fundamentally an exercise in *probabilistic estimation*, offering a set of strong priors that must be ruthlessly updated as empirical reality delivers new data. As the neurotechnology landscape matures and the fog clears, we must remain relentlessly willing to **re-**

⁶⁹. Goodhart's law: when a measure becomes a target, it ceases to be a good measure; optimize too hard for a proxy and it stops tracking the thing you cared about.

assess our trajectory and critically question whether neurotechnology *remains a net-positive* technology to continue developing.

Neurotechnologies are not a panacea. Even a flawless, zero-latency system capable of reading and writing a human mind with perfect fidelity still leaves us with an immense, daunting mountain of alignment, ethical, and architectural problems to solve. It is only *one piece* of a vastly larger puzzle.

However, if we ground ourselves in first principles and project out into the long-term horizon, say the year 2060, **BCI will be a prerequisite for humanity to remain consequential.** Once machines operate at a scale a million times more capable than biological humanity, able to execute, reason, and iterate better than us in every conceivable domain, establishing a robust, high-bandwidth control and interaction surface will be an **absolute necessity.**

24 The Decade Ahead

“The future is already here; it’s just not very evenly distributed.” — William Gibson

We tend to look at humanity’s greatest achievements as singular, explosive events: the day the bomb dropped, the day the internet turned on, the day the first LLM went viral. But hindsight is a terrible historian, always eager to dramatize and compress events. In reality, deep technological shifts happen more *like water freezing*: a long, invisible drop in temperature, then the first sudden slush, and then a thousand small crystals locking into place until, seemingly all at once, the ground beneath you is solid ice.

The next ten years of neurotech will follow the same narrative, the same one we just witnessed with language models: a long quiet *under-the-radar* phase, then *sudden visibility*, followed by *ubiquity*, and finally the *strangeness* that comes with a new world order.

We now turn to the future. The next ten years will likely be the most transformative in human history. **AI will give machines intelligence, robotics will give them bodies, and BCI will begin connecting them directly to the human mind.** The Old World is still standing, but the scaffolding of the new one is already everywhere.

2026–2028: First slowly

The next 24 months will look the way 2018 to 2020 looked in AI from the inside. *Calm, but accelerating.*

The first usable **foundation models for brain decoding** will ship publicly, trained on EEG, ultrasound or MEG. Frontier AI labs will start exploring the use of neural data. At the same time, the first generation of non-invasive consumer devices a healthy adult would actually choose to wear will begin shipping at scale, somewhere between sub-vocal text decoding and continuous attention monitoring.

BCI will leave its purely academic, medical niche and start bleeding into the broader culture. You hear friends bring it up at dinner, a high-production documentary drops, a YouTube deep-dive goes viral. It will stop feeling like a sci-fi trope and start feeling more like an impending reality.

As the tech matures, impressive lab results begin to stack up. Teams demonstrate the ability to decode complex language, reconstruct visual images from the visual cortex, and map fine-grained user preferences directly from neural signals. This compounding momentum changes how capital and talent move, top-tier software engineers and AI researchers who were entirely focused on LLMs start pivoting their careers toward the human machine interface.

The gravitational pull of BCI is growing.

Crucially, this technology accelerates a shift that AI already started, the human role continues to drift away from being an *operator*, and toward being *a source of judgment, taste, priority, and intent*.

By early 2028, a frontier AI lab will announce a neural-data partnership, explicitly framing it as a training modality or alignment investment. Shortly after, the first landmark peer-reviewed paper will drop, showing that brain-tuning data directly improves a flagship model's reasoning or alignment metrics.

By the end of 2028, the calm phase ends. It will become blindingly obvious to anyone paying attention that BCI is moving much faster than the public consensus expected. *The dam will begin to break.*

2028–2031: And then suddenly

The late 2020s and early 2030s will look the way 2022 to 2025 looked in AI. *Loud.*

The first BCI primitive will hit 10 million active users. Sub-vocal text will begin displacing voice commands and physical typing for short interactions. Some people will start reading mail, sending messages, and dictating notes entirely in their heads, the same way mobile typing displaced laptop typing a decade earlier. BCI will also become a more central tool for alignment, and the first peer-reviewed demonstrations that brain-tuning improves frontier-model metrics on standardized benchmarks will happen around 2028 or 2029. This is the era where people stop calling neurotechnology an “*emerging*” field. By 2031, BCI will hit its *Tesla moment*: a flagship consumer product that the general public can name will appear, making the technology mainstream.

As this transition happens, the hardware will iteratively vanish from daily life. The form factor will shrink into earbuds, a beanie, the temples of AR glasses, a hat or a simple woven headband. It becomes wearable all day, practically invisible, and socially acceptable.

BCI will start embedding itself deeper into the social fabric. People begin using it more regularly at work and in social settings. It shows up more often in movies, comes up more often in politics, and public figures start talking about it.

As BCI improves, it shifts from decoding coarse commands to decoding **fine-grained intent**. You no longer have to formulate a sentence to trigger an action, it becomes possible to interact with AI using only your thoughts, and starts to become more effective than typing or speaking. **Thought-to-text** and **thought-to-speech** become familiar concepts inside AI labs, and “*how do we make LLMs easier to interface with through the brain?*” becomes a hot topic.

This unlocks capabilities that sound genuinely absurd by today’s standards. You think of a query, and your earbuds whisper the answer back. You beam a raw concept or a melody directly to a friend’s headset. You record memories, dreams or thoughts and play them back as images or videos. And with the rise of embodied AI and robotics you can even look at a domestic robot, think about a task, and watch the machine execute it with near zero friction.

At this stage, a reasonable number of knowledge workers will have tried a non-invasive headset, much like VR today, except this time the technology *actually works*. Once people can touch it, wear it, and feel the interaction for themselves, the technology becomes tangible, shifting the public conversation from metaphysical speculation to an immediate reality. The public conversation will get more vivid overnight.

Simultaneously, the invasive side of the market will mature rapidly. Implants will scale from hundreds of clinical-trial patients to tens of thousands of medical users, nearing a hundred thousand by 2031.

Crucially, for invasive BCI, the delivery mechanisms will become far less barbaric as the field begins moving beyond the “skull drill”⁷⁰. The industry will start moving away from open-brain surgery toward minimally invasive delivery. Companies like Synchron will deploy stents up the jugular vein⁷¹ to read the motor cortex from the inside of a blood vessel, avoiding the skull entirely. Competitors like Neuralink will advance techniques that avoid piercing the dura mater⁷², while others push toward flexible bioelectronics and injectable neural dust: tiny implanted sensors designed to record neural activity with less surgical trauma.⁷³

These years won’t just be about surreal technology; they will also be ruthlessly competitive. BCI will transition from a fun gadget to an economically necessary tool, and governments, massive corporations, and institutional capital will race to secure position. As the stakes rise, mental privacy will become a mainstream political fight, showing up on social media, in regulation, and in arguments at home. **Neurorights** will become normal policy language.

⁷⁰. For most of its history, serious invasive BCI meant opening the skull and placing electrodes on or in the brain. That may still produce the highest-fidelity signals, but it is a brutal onboarding flow for a technology that hopes to scale.

⁷¹. As much as this sounds like a medieval torture technique, compared with opening the skull, going up the jugular vein is the civilized option: well understood and low risk. (Out of context, this is probably the kind of sentence which might sound slightly awkward in court)

⁷². Neuralink’s architectural updates focus on deploying flexible threads without requiring a full breach of the protective dura mater, radically reducing surgical risk and recovery time.

⁷³. Neural dust: millimeter- and eventually micrometer-scale wireless sensors placed near neurons, powered and read out by ultrasound. The appeal is to be able to record from inside the brain without wires or bulky implants.

By 2029, cumulative venture funding in BCI will cross the \$8 to \$10 billion mark. Sub-vocal text will become a viable input for AI glasses, and the first wave of non-invasive consumer products will begin shipping at scale.

By 2030, the industry will cross the 10-million shipped units threshold. This milestone triggers the arrival of a US federal neural-data framework with actual enforcement teeth, alongside the first frontier AI model cards explicitly disclosing neural data in their training mix.

That same year, the medical sector will hit 10,000 active invasive BCI implants globally, demonstrating the safety and scalability of surgical deployments.

By 2031, the true *Tesla moment* arrives. Cognitive-state-aware AI assistants are firmly on a path to become the dominant, default interface to machines, permanently changing how humans interact with digital infrastructure, setting humanity on a path toward a full merge by the 2040s.



2036, a Tuesday morning.

2031–2036: Ubiquity

By the early 2030s, BCI crosses the threshold into true ubiquity, it is no longer a technology you have to explain or justify. Everyone has an opinion on it, everyone sees its footprint on society, and the initial debate over whether it is real dissolves entirely.

The most profound shift of this era is the evolution from prompting an AI to *thinking with* an AI. Collaboration turns into symbiosis, the days of pulling out a phone to type a query start to look ancient. Instead, the AI runs as a continuous background process alongside your mind. The boundary between your biological memory and the machine's data retrieval begins to blur.

This creates a true **cognitive exoskeleton**. Just as the smartphone turned humanity into a primitive, high-latency cyborg network in the 2010s, AI and BCI create a low-latency, always-on **exocortex**. It amplifies human capability at a deeper, more structural level than any screen ever could.

This transformation does not hit everyone at once. The future remains unevenly distributed, and a large portion of the population continues to work and live using the baseline tools of the mid-2020s. But for an *ever-increasing*, hyper-productive segment of the global workforce, operating without this cognitive layer feels like trying to run a modern business without electricity.

At the same time, computing transitions from *episodic to ambient*. The headset or implant reads your cognitive state continuously, not just when you consciously issue a command. AI agents monitor your cognitive load, attention levels, and mental fatigue in real time. If a complex technical task spikes your stress, the system can automatically dial back distractions, curate your digital layout, and surface the exact reference material you need before you even realize you are stuck. Prompting becomes a relic of the early AI era; ambient computing becomes the new default.

This is also the moment where the “**write**” side of neurotech starts picking up real commercial momentum. Up until now, the field was heavily dominated by high-fidelity reading. By the middle of the decade, precision neuromodulation procedures begin moving into the consumer mainstream, becoming as harmless as LASIK surgery. It is now possible to write precise thoughts and feelings to the brain. By 2037, the first augmented adults will think measurably faster than their own unaugmented selves.

Ultimately, the early 2030s is the point where four massive, independent technologies collide. High-bandwidth non-invasive BCI, frontier artificial intelligence, large-scale humanoid robotics, and early brain-writing capabilities. All at once, they stop being separate industries. They fuse into a **single, compounding tech stack**. Mirroring the way smartphones, the cloud, mobile broadband, and app stores combined in the early 2010s to create the modern digital world.

Over time, read becomes read-and-write, industrial and humanoid robotics systems adopt BCI as their natural control interface, the bandwidth bottleneck between human intent and machine execution closes in earnest, and the broader AI alignment problem makes significant progress, in large part thanks to advances in neurotechnology.

By 2032, direct robotic control through BCI becomes a highly credible, multi-billion-dollar enterprise category. The same year, the first heavily capitalized neurotech conglomerate begins offering fully bidirectional, consumer-grade read-and-write systems designed specifically for healthy adults.

The **gravitational pull of BCI** is now second only to a few technologies; alongside AI and robotics, it *redefines* how the world operates.

2036 onwards: The strange phase

After 2036, the timeline loses its linear feel. We cross a civilizational horizon where the boundary between human and machine disappears.

The catalyst is the *Write* revolution. For 20 years, neurotech was nearly exclusively about listening to the brain. Now, we talk back. The industry figures out how to insert rich sensory data, structured concepts, and motor skills directly into the cortex. The *Matrix* skill upload⁷⁴ transitions from a sci-fi trope to an emerging feature of reality.

On a daily basis, this manifests as a fully **programmable state of mind**. You gain granular control over your own neurochemistry. You tap a screen, or just issue an internal intent, to induce instant, restorative sleep, mute an adrenaline spike before a presentation, or lock into a flawless four-hour hyper-focus state without pharmaceuticals.

But as the technical barriers fall, the philosophical ones become *existential*. When you can rewrite a mental state at will, the concept of the *self* fractures. If a machine can dial down your anxiety or inject a memory, where do the boundaries of your original identity end? Humanity is forced to invent entirely new legal and moral categories as the mind is no longer an unmappable black box.

We are now navigating *dangerous, uncharted territories*.

This convergence enables true **proto-telepathy**. For millennia, human collaboration has been bottlenecked by language, an incredibly low-bandwidth, *lossy compression* format for thought.

⁷⁴. The *Matrix* skill upload: the scene where Neo has kung fu loaded straight into his brain and wakes up knowing it. Great movie; do me a favor and go check out the [clip](#).



Proto-telepathy: A whole idea transmitted without being translated into words.

With bidirectional communication, we unlock **conceptual sharing**. Instead of spending an hour translating a complex spatial map or nuanced feeling into words, you transmit the raw, multi-dimensional concept directly to a collaborator's mind. They do not just hear your explanation; they grasp the mental model *exactly* as it exists in your head.

For better or worse, emotions themselves become transmissible, and real-time **transmission of emotional states** challenges the very nature of human isolation. When you can *feel* the precise texture of another person's grief or fear, human conflict and relationship dynamics undergo a brutal structural rewrite.

As the years roll forward, the loop between mind and machine tightens into a complete symbiosis. The baseline definition of the impossible keeps shifting, leaving the twentieth-century world feeling like a distant memory.

Maya in 2038

Maya wakes up. The subtle neural node behind her ear already knows she slept poorly; her morning agent has rescheduled her 9:00 AM meeting before she has even opened her eyes.

She forms the internal intent for *coffee*, and the kitchen downstairs hums to life.

The first hour of work used to be emails. Now it is an unvoiced conversation. She thinks toward her assistant, the assistant thinks back, and the rhythm of the exchange runs at the speed of inner thought. The decoding model does not read the words she would have typed, but the *raw, uncompressed* shape of what she thinks.

When her attention drifts toward a lingering problem from yesterday, the assistant seamlessly surfaces the relevant context without being asked. When a drafted response from the AI is slightly off, Maya's recognition of the wrongness reaches the feedback loop milliseconds before her physical hands could ever move. She corrects the model in real time, without ever articulating the correction.

When she needs to recall a specific metric or look up a complex financial projection, she does not feel like she is searching an external database. She simply feels like she suddenly remembers the answer, it is seamlessly injected into her mind. The boundary between "a human thought" and "a machine thought" within any given reasoning has *blurred*.

By lunchtime, Maya has cleared a volume of creative and analytical work that, in 2026, would have taken her more than a week. But it does not feel special at all, it just feels like a normal Tuesday.

The infrastructure that made all of this ordinary arrived in increments. For years, the typical knowledge worker wore a non-invasive headset every single day. By 2035, the surgical and regulatory floor for elective, low-risk neural modules dropped to where laser eye surgery sat a decade earlier. What began as medical restoration for late-stage ALS and Parkinson's patients slowly became a **standard surgical upgrade**. Maya got hers last year, right after her close friend did.

Now, she no longer remembers what it felt like to have to compress every thought through her thumbs. The small, bidirectional module interfacing directly with her cortex is now an inherent part of her life and her *self*. It even helps with memory consolidation overnight; she sleeps deeper and dreams more vividly.

Late in the afternoon, Maya steps away from her digital flow to paint. She picked up the hobby recently, though the learning process looked nothing like it would have a decade ago. She did not spend months failing at perspective or color theory. Instead, she *acquired* the foundational muscle memory and spatial understanding directly through her neural node. As she waits for a layer of oil to dry, her mind drifts to a dense philosophy book on visual aesthetics. She never actually sat down to read it, she simply accepted the **conceptual transfer** over the weekend, and by morning, the arguments were entirely legible to her. As

she blends a shadow on the canvas, she casually executes a technique she never consciously practiced, guided by the vivid memory of a text she never opened.

Sitting back, Maya spots a scattering of cups and streamers in the garden from her children's birthday party yesterday. With a fleeting pulse of intent, she directs a small fleet of domestic robots to clear the lawn. The machines glide into motion as she slips seamlessly back into her afternoon flow, lost in her own thoughts.

Later, she enters the design phase of a new architectural project; she needs to finish the first iteration by tonight. She does not write code or draft blueprints, software and structure have stopped being typed; they are simply imagined into existence. She envisions the layout, materiality, and light of a new building with the fluid, intuitive ease of a child building in Minecraft⁷⁵.

By early evening, when her partner walks into the room, Maya does not need to ask how his day went. She can literally *feel* it; she instantly catches the raw texture of his emotional state, the lingering friction of a difficult morning dissolving into relief as he sits down. The neural link creates a silent, high-bandwidth empathy, allowing them to share the exact form of a feeling without relying on the lossy compression of words.

Lately, the technology has pushed into territory that used to be **strictly metaphysical**. Last night, her friend, a fervent alpinist, streamed an execution of her ascent of Mount Everest to Maya's son, who loves mountaineering. He did not just watch a video; he sensorially teleported. Through the shared neural experience, he felt the crisp altitude, the biting cold, and the sharp awe of looking down from the ridge. Memory, and even emotion, have stopped being constrained to first-person acquisition.

75. For those living in a cave: Minecraft is the best-selling video game of all time, built around placing blocks to construct anything you can imagine.

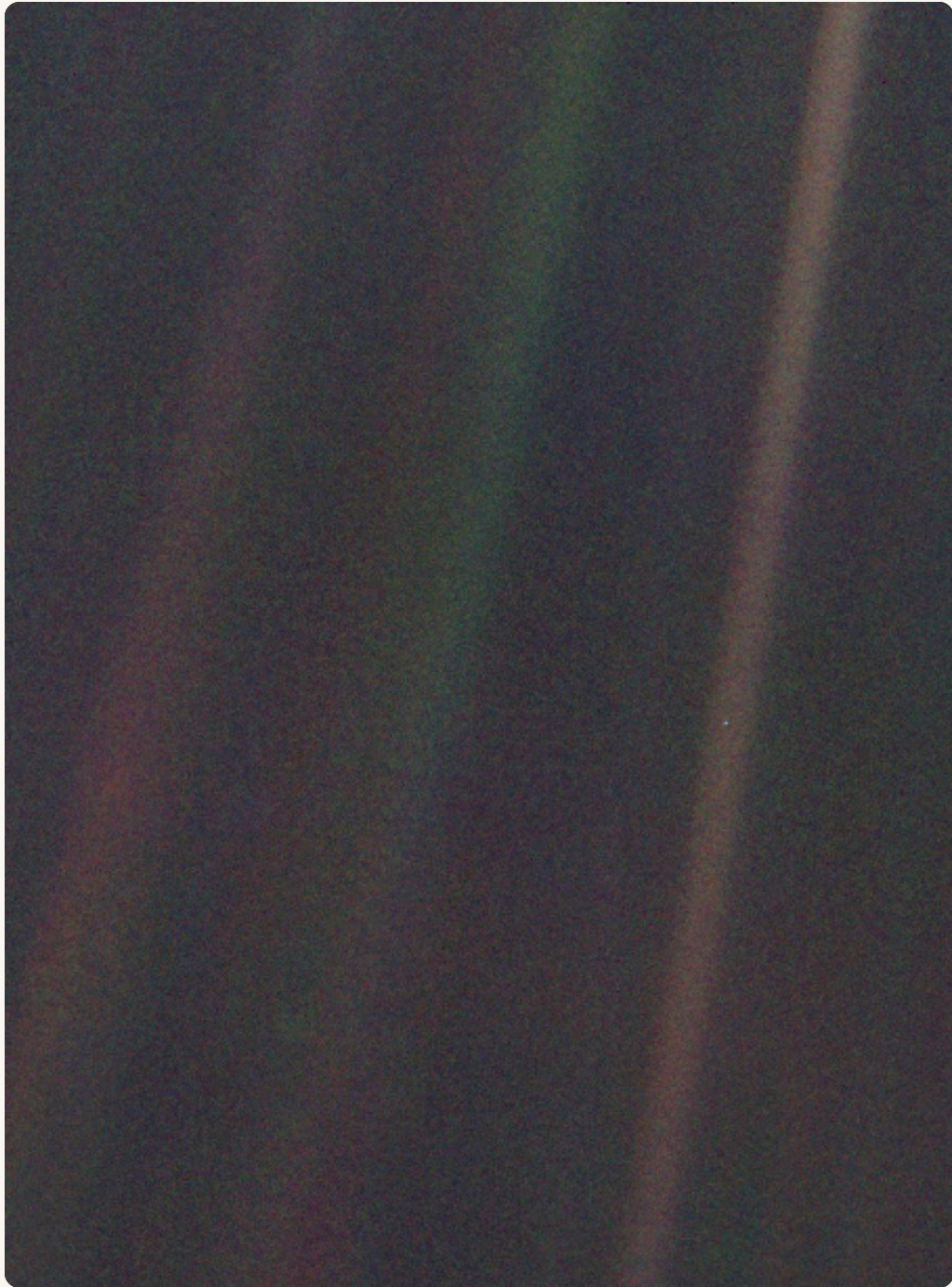


A summit you never climbed, remembered as if you had.

Even her own past has become fully legible. On her break, Maya occasionally replays her own childhood memories at full resolution; the full experience decoded from the original neural trace and rendered back into her *sensorium*. She can relive a specific morning from her childhood, or re-experience a conversation with someone who is long gone, catching the exact texture of their voice.

Only in retrospect did it become clear that a threshold had been crossed. While it was happening, every advance looked incremental, and every adoption looked like another ordinary personal choice. There was no master switch, and no dramatic Thursday when everyone plugged into the Matrix.

The future took hold quietly, as it usually does. First the device was *a curiosity*, then *the norm*, then it became *strange to live without*, and finally it *vanished altogether from attention*, like electricity or the internet a few decades prior.



A Busy Mote of Dust: Earth, captured from six billion kilometers out by Voyager 1 in 1990.

25 Conclusion

For the first time in the history of the known universe, a general intelligence other than our own has emerged. Our long solitude is finally ending. *Matter arranged by human hands has begun to think.*

With this comes the **greatest challenge** in the history of our species: artificial intelligence. It forces us to answer impossible questions on an engineering timeline: *What is being human when the boundary of the mind becomes permeable? What should we become now that the future is an explicit design choice? And what is our role in the universe that follows?*

Reality is bending, and AI is *distorting* our society faster than we can adapt. We are entering a world governed by systems we did not evolve to understand, operating at speeds we were never built to match. And whether you welcome it or dread it, this new age is already upon us.

But its trajectory is not yet fixed. The coming decade is our remaining chance to shape the future before AI capabilities move beyond our reach, before the Phase Lock Window closes *permanently*.

This manifesto began with a mismatch: machine intelligence is scaling faster than the human capacity to oversee it. And as this bandwidth gap is widening, BCI has emerged as one of the few technologies capable of addressing it at its roots.

As of 2026, BCI is still nascent and lagging behind AI, but the path forward is becoming legible: better hardware, larger neural datasets, and bigger models are bringing within reach the *breakthroughs* needed to move neurotechnology beyond the lab. And as BCI matures, its convergence with AI will shape the trajectory of our civilization.

Ultimately, **AI will scale BCI, and BCI will scale AI.**

BCI will not be sufficient on its own to solve the numerous challenges we face, but through *alignment-by-integration*, it will be an important and necessary part of any durable solution. As *alignment-by-control* becomes ever more fragile, it will create a channel for oversight that scales with the intelligence it supervises, and, in time, the foundation of a symbiosis between biological and artificial minds. In a world increasingly shaped by AI, it tilts the balance of power back toward human intent, ensuring that the systems we build remain instruments of human consciousness and serve the flourishing of life.

But the future will not be prosperous by default, we will have to *build it deliberately*. And the task of forging it does not belong to some abstract, distant generation; it falls upon *all of us*, right now, and above all on those aware enough to see what is coming.

If you are reading this, *you are one of them*.

Find where you can push, and *push*. Building for the Old World is mostly a waste of time at this point; all that truly matters now is ensuring humanity can prosper in the age of AI. The work to be done extends beyond BCI⁷⁶, and I encourage **you** to figure out how you can take part in shaping this future. Whether you are a builder, an investor or simply someone who cares, feel free to reach out to me with

any thoughts and ideas; and if this manifesto has changed how you see the next decade, *put it in front of someone* who should be helping shape it.

AI will be humanity's last prowess, the final expression of our ingenuity before machines take over the instruments of creation. Neurotechnology will help ensure that as they do, humanity remains a composer of what comes next, rather than a fading echo within it.

As Humanity and AI enter **Phase Lock**, the choices we make over the next decade will define the *centuries* to come.

The path ahead remains unmarked, the map unfinished, and the horizon unlike anything our species has seen before.

But still, we advance, wading through the fog, drawn by the light of an unwritten tomorrow.



The road ahead: Wading through the fog, drawn by the light of an unwritten tomorrow.

76. [Here](#) is a concise but complete mental model of what the AGI transition encompasses. You can like the Tweet to be added to a group chat with fellow readers who had the courage to complete this manifesto.

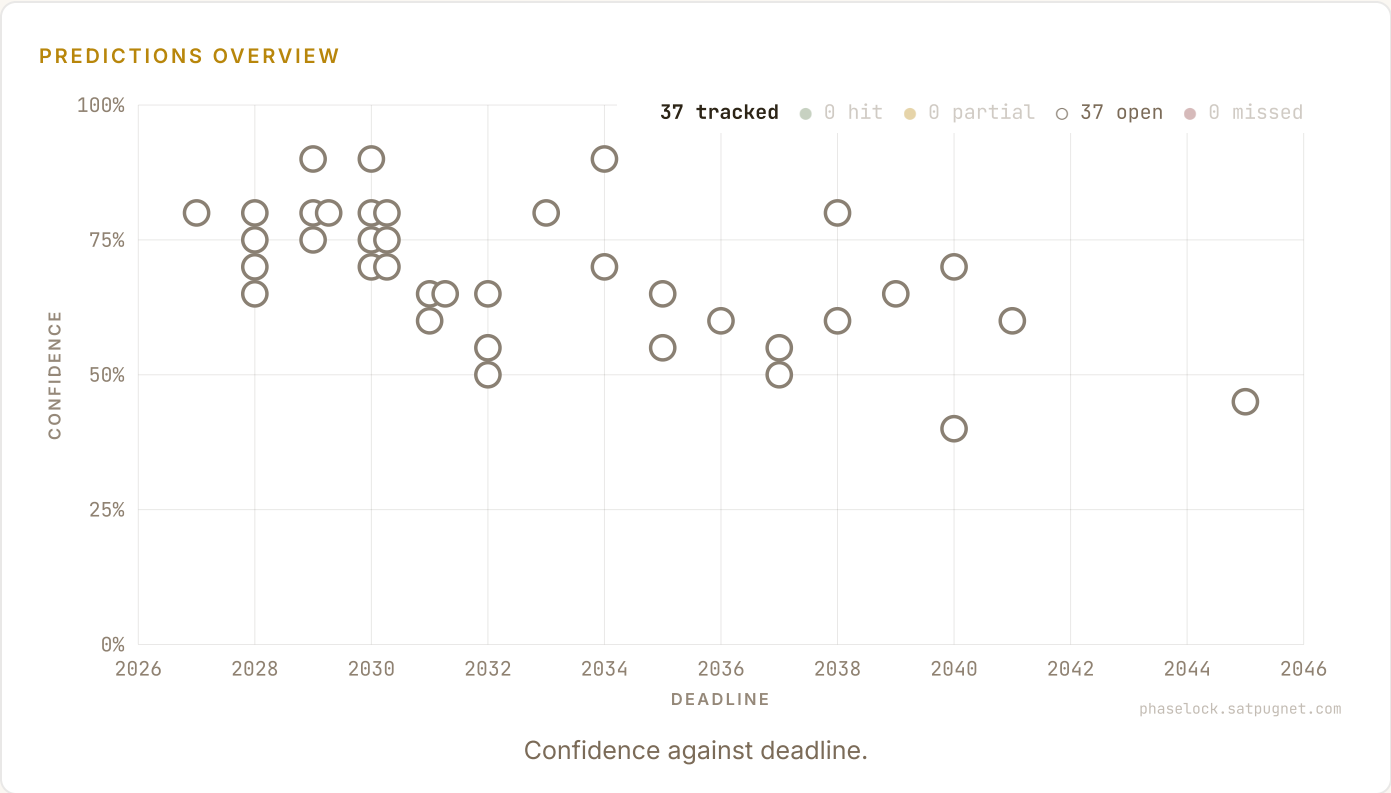


APPENDIX

Predictions, acknowledgments, resources, and the modality map.

Appendix A: Predictions to Revisit

I made a number of predictions throughout this manifesto. This appendix collects the main ones in one place, each with a target year and a confidence, grouped by theme. I will update them as reality unfolds.



2027

- A non-invasive BCI demonstrates reliable intention decoding that generalizes across unseen users

Decoding · 80%

2028

- A frontier AI lab publishes a paper or blog post identifying BCI as an alignment lever

Alignment · 80%
- A frontier lab publishes evidence that brain-tuning measurably improves a flagship model on a standardized benchmark

Alignment · 75%

- The first non-invasive consumer BCI allowing direct communication with AI launches Decoding · 70%
- AI labs begin incorporating neural data into frontier-model training mixtures Data · 65%

2029

- Cumulative venture funding for BCI exceeds \$8 billion Capital · 90%
- A frontier AI lab makes an eight-figure investment in BCI Capital · 80%
- A non-invasive BCI reliably decodes inner speech at over 120 words per minute without using muscle signals Decoding · 80%
- A non-invasive BCI reaches useful performance on a new user after less than five minutes of calibration Decoding · 75%

2030

- The United States designates BCI a critical and emerging technology Governance · 90%
- An open-source brain foundation model exceeding one trillion parameters is released and used by an independent team Data · 80%
- The number of active invasive BCI implants exceeds 10,000 worldwide Adoption · 80%
- A neural dataset reaches 10 million hours of recorded brain activity Data · 75%
- Cumulative shipments of non-invasive consumer BCI devices exceed 10 million units worldwide Adoption · 75%
- A major AI lab publicly uses neural feedback in a post-training or model-evaluation pipeline Alignment · 70%
- The United States enacts a federal neural-data framework with enforceable limits on collection, resale, or secondary inference Governance · 70%

2031

- A consumer BCI reconstructs a user's mind's eye as an image on their phone Decoding · 65%
- BCI has its "Tesla moment": a flagship consumer product becomes widely recognizable Adoption · 65%

- A non-invasive BCI integrated into glasses, earbuds, or a headband supports all-day use and meaningful decoding in ordinary environments Hardware · 60%

2032

- A major neurotechnology company faces regulatory action or mass litigation over neural-data misuse Governance · 65%
- A commercial system enables direct neural control of robots or drones without a manual controller Interface · 55%
- A venture-backed company ships a bidirectional read-and-write BCI for healthy adults Neural Write · 50%

2033

- The first brain-to-brain transfer of a complex, nonverbal concept between two people is demonstrated and independently verified Decoding · 80%

2034

- BCI systems restore vision, hearing, and limb control to near-natural levels Neural Write · 90%
- A BCI decodes imagined scenes across multiple sensorial modalities, including sound, motion, and felt sensation Decoding · 70%

2035

- Cognitive-state-aware BCI is used by more than 5% of frontier-model users Adoption · 65%
- An elective BCI for healthy adults can be implanted through a minimally invasive outpatient procedure without open craniotomy Hardware · 55%

2036

- A BCI reliably induces a targeted mental state, such as sleep, focus, calm, or alertness, on demand Neural Write · 60%

2037

- A neural implant or neuromodulation system measurably raises a healthy adult's reasoning speed or working memory beyond their own unaugmented baseline on standardized cognitive

tests

Neural Write · 55%

- A BCI transfers a motor skill or structured concept in a way that measurably improves performance without conventional practice Neural Write · 50%

2038

- A BCI reconstructs and replays a specific autobiographical memory with recognizable detail Neural Write · 80%
- A recorded multimodal experience from one person can be neurally replayed to another with recognizable sensory and emotional content Neural Write · 60%

2039

- A BCI writes or alters the content of an autobiographical memory Neural Write · 65%

2040

- Active invasive BCI use exceeds ten million people worldwide Adoption · 70%
- A BCI reads and writes neural activity at neuron-level precision across the whole brain, closing a full bidirectional cognition channel Neural Write · 40%

2041

- A credible whole brain emulation of a specific human is demonstrated Emulation · 60%

2045

- Multiple humans and an AI collaborate through a shared neural workspace, exchanging concepts without language Interface · 45%

Appendix B: Friends and Thought Partners

This manifesto did not come to life in a vacuum. The people below have shaped how I see this field over years of conversation. Each, at some point, moved the thinking forward. The smart claims are theirs; the errors and overclaims are mine.

One person holds a special position here: my partner Olya, who read every draft, caught what did not work, and refused to let this sit in my backlog for another two years. It exists in finished form thanks to her.

Friends and Thought Partners

On the hardware side: [Sumner Norman](#) and the [Forest Neurotech](#) / [Merge](#) crew, [Mikhail Shapiro](#) at Caltech, [Ryan Field](#) at [Kernel](#), [Nishita Deka](#) at [Sonera](#), [Christina Maher](#), [Lev Chizhov](#), [Jay Sanguinetti](#) at [Sanmai](#), [Mehmet Günal](#) at [Brill Neurotech](#), [Hon Weng Chong](#) at [Cortical Labs](#).

On data and decoding: [Trent McConaghy](#) at [Psyche](#), [Jonathan Xu](#) and the [AllJoined](#) team, [Andreas Tolias](#) and [Sophia Sanborn](#) at [Enigma](#), [Sharena Rice](#) at [Sanmai](#), [Milan Cvitkovic](#) at [Integral Neuroscience](#), [Avery Krieger](#) at [Constellation](#), [Naomi Bashkansky](#) at [OpenAI/Conduit](#).

On consumer-facing devices: [Ramses Alcaide](#) at [Neurable](#), [Andreas Forsland](#) at [Cognixion](#), [Meron Gribetz](#) at [Inner Cosmos](#), [Akshat Sharma](#) at [Orbit](#), [Peter Schlecht](#) at [Braingrade](#).

On the invasive and endovascular frontier: [Matt Angle](#) and [Vikash Gilja](#) at [Paradromics](#), [Tom Oxley](#) at [Synchron](#), [Takufumi Yanagisawa](#) at [ivec](#), [Marc-Joseph Antonini](#) at [Neurobionics](#).

On whole brain emulation: [Isaak Freeman](#) at [MIT/Capable](#), [Catalin Mitelut](#) and [Christian Larsen](#) at [Netholabs](#), [Daniel Burger](#) at [Eightsix Science](#), and [Adam Marblestone](#) at [Convergent Research](#).

On AI alignment and NeuroAI: [Seth Bannon](#) at [50y Ventures](#), [Mike McCormick](#) at [Halcyon Futures](#), [Judd Rosenblatt](#) and [Diogo de Lucena](#) at [AE Studio](#), [James Fickel](#), [Patrick Mineault](#) at the [Amaranth Foundation](#), [Matthieu Kratz](#) at [Coefficient Giving](#), [Frank Barbosa Nakasako](#) at [Palaestra Research](#), [Evan Miyazono](#) at [Atlas Computing](#), [Sean Escola](#) at [Inductive](#), and [Ryota Kanai](#) at [Araya](#).

On funding and infrastructure: [Michael Andregg](#) and [Maximilian Schons](#) at [Eon](#), [Raphael Certain](#) at [Clarity](#), [David Langer](#) at [Lionheart](#), [Peter Zhegin](#) at [e184](#), [Olha Dolinska](#) at [Oxford](#), [Hilmar Eggertsson](#) at [6 degrees capital](#), and [Ryan Zurrer](#) at [Dialectic](#).

On conveners and community: [Diana Saville](#) at [BrainMind](#) and [Allison Duettmann](#) at the [Foresight Institute](#), who bring much of this field together.

On the outside view: [Karim Beguir](#) at [Instadeep](#) and [Raphael Volpert](#) at [Alexandria](#), [Yash Shevde](#) at [Ursa Bio](#), [Ansh Chopra](#) at [ZeroBillion](#), [Viktor Bowallius](#) at [ThinkIF Technology](#), [William Knottenbelt](#) at [Imperial College London](#), [Adrian Johnston](#) at [Elyos AI](#), [Mate Polat](#) at [Varto AI](#), [Atlal Boudir](#) at [Louise](#), [Daniel Faggella](#) at [Emerj](#), [Rodolfo Rosini](#) at [Vaire Computing](#), [Nikolas Ioannou](#) at [Hyper](#), [Dan Girshovich](#) at [Tools for Humanity](#), and [Auguste Pugnet](#), who do not build in neurotech but have been valuable thought partners nonetheless.

Appendix C: Further Resources

I mentioned a lot of companies and labs throughout the essay. I am consolidating the ones worth keeping an eye on here, for anyone wanting a bird's-eye view of the field.

Companies and Labs Worth Watching

Non-invasive hardware: [Cerca Magnetics](#) (OP-MEG), [Kernel](#) (fNIRS), [Nudge](#) (ultrasound), [Butterfly Network](#) (ultrasound chip), [Aleph Neuro](#) (ultrasound), [Sonera](#) (biomagnetic neural sensing).

Non-invasive data and decoding: [AllJoined](#), [AE Studio](#), [Mind Co](#), [Neurable](#), [Brill Neurotech](#), [Wisear](#), [Orbit](#), [Conduit](#), [Pyramidal](#), [Constellation](#), [Sabi](#).

Neural data for AI: [Inductive](#).

Open hardware and community: [OpenBCI](#).

Sub-vocal and peripheral neural interface: [CTRL-labs \(now Meta neural interface\)](#): the EMG sub-vocal/intent path that is the closest non-invasive cousin to BCI in the consumer interface frame.

Prior consumer non-invasive wave (mostly cautionary): [Muse](#), [Emotiv](#), [NextMind](#) (acquired and then wound down by Snap).

Neuromodulation and stimulation: [Sanmai Technologies](#), [META-BRAIN](#) (EU magnetoelectric + ultrasound neuromodulation), [Integral Neuroscience](#).

Medical-grade non-invasive: [NeuroPace](#).

Minimally invasive and invasive: [Paradromics](#), [Synchron](#), [Neuralink](#), [Precision Neuroscience](#), [Starfish Neuroscience](#), [Neurobionics](#).

Adjacent and meta: [Psyche](#), [Forest Neurotech](#) (now [Merge](#)), [Inner Cosmos](#), [Cognixion](#), [Enigma](#), [Lymbic AI](#), [Clarity](#), [Cortical Labs](#).

Adjacent, endovascular, and emerging: [ivec](#), [Araya Inc.](#), [Neuradaptive](#), [Braingrade](#), [Synlinx](#) (China's national BCI incubator, Shanghai).

Whole brain emulation: [Eon](#), [Eightsix Science](#), [Netholabs](#).

Investors and supporting capital: [Amaranth Foundation](#), [Jumpspace Ventures](#), [Lionheart Ventures](#), [e184](#).

In addition to companies, here is a list of write-ups for anyone wanting to go deeper into neurotech:

Further Reading

Companion essays

- By myself, [Humanity's Last Prowess](#). My overarching framework of what matters for the AGI transition.
- By myself, [The Future in One List](#). A broader map of the technology that will define the coming decade.

Related to the essay

- Leopold Aschenbrenner, [Situational Awareness](#), 2024. An AGI Manifesto.
- [AI 2027](#), 2025. A scenario forecast for AI over the coming years.
- Sam Altman, [The Merge](#), 2017. One of the early arguments that human and machine cognition would converge.
- Rich Sutton, [The Bitter Lesson](#). Two pages. Referenced several times in this essay.

Onboarding to AI alignment

- **Best single starting point:** Robert Miles, [Intro to AI Safety, Remastered](#). The clearest one-video primer on why this matters at all.
- For more depth, Rob Miles's [full AI safety playlist](#), the [AI Safety Map](#) of the field's organizations and research agendas, and [aisafety.info](#) for a curated Q&A-style guide.

AI scaling laws

- Kaplan et al., [Scaling Laws for Neural Language Models](#), OpenAI, 2020. The paper that turned “more data and compute keep helping” from an intuition into a power-law. The empirical foundation under the whole Bitter Lesson argument.
- Hoffmann et al., [Training Compute-Optimal Large Language Models](#) (Chinchilla), DeepMind, 2022. Corrected the data-to-parameters ratio modern frontier models train at.

Core BCI evidence

- Tang et al., [Semantic reconstruction of continuous language from non-invasive brain recordings](#), Nature Neuroscience, 2023.
- Willett et al., [High-performance brain-to-text communication via handwriting](#), Nature, 2021.
- Metzger et al., [A high-performance neuroprosthesis for speech decoding](#), Nature, 2023.
- Scotti et al., [MindEye: fMRI-to-image with contrastive learning](#), 2023.

Foundation models for the brain

- Caro et al., [BrainLM: A foundation model for brain activity recordings](#).
- Banville, d'Ascoli, Dahan, Rapin et al., Meta AI, [NeuralBench: A Unifying Framework to Benchmark NeuroAI Models](#), 2026.

NeuroAI and alignment

- Trent McConaghy, [bci/acc: A Path to Balance AI Superintelligence](#) ▶. The case for BCI as an alignment lever, from inside the field.
- Patrick Mineault and Joanne Peng, [NeuroAI for AI Safety](#), Amaranth Foundation. A rigorous roadmap for neuro AI safety.

Community resources

- The community-curated [awesome-brain-decoding](#) repository.
- [Brain-Computer Interfaces: Exploring the Last Frontier](#), Martin Krag, byFounders. A thoughtful investment thesis presenting a closely aligned vision.

Adjacent: brain emulation and connectomics

- Anders Sandberg and Nick Bostrom, [Whole Brain Emulation: A Roadmap](#), Future of Humanity Institute, 2008. The canonical reference for what WBE would actually require, from substrate scanning through compute.
- Isaak Freeman, [WBE feasibility thesis](#). Recent cost analysis putting the all-in figure for human emulation at roughly \$5-50 billion.

Appendix D: The Modality Map

A reference breakdown of the modalities named throughout this manifesto. The interactive comparison is in [the modality map](#) ▶; below is each modality in detail:

The Full Modality Map

Non-invasive BCIs

EEG: Measures electrical activity at the scalp with millisecond precision. Cheap and scalable but spatially noisy, it is the main test bed for early brain foundation models like [AllJoined](#)'s.

fNIRS: Uses near-infrared light to track blood oxygenation, an indirect signal that lags neural activity by several seconds. It is wearable, and [Kernel](#) is one of its leading commercial platforms.

Ultrasound: Uses sound waves to image blood flow and stimulate targeted regions deep in the brain. Still early, but promising for compact read-and-write interfaces. [Butterfly Network](#) put ultrasound on a chip; [Nudge](#) and [Merge Labs](#) are applying it to the brain.

MEG: Measures magnetic fields produced by neural activity with millisecond precision. Once limited to room-sized cryogenic scanners, it is becoming wearable through OPMs.

fMRI: Uses powerful magnetic fields to track changes in blood oxygenation, an indirect measure of neural activity. It offers millimeter-scale spatial resolution and has enabled some of the best non-invasive [semantic reconstructions](#), but it is slow, expensive, and confined to room-sized scanners.

Invasive BCIs

Endovascular: Guides an electrode through the jugular vein into a blood vessel beside the motor cortex. [Synchron](#)'s Stentrode avoids open-brain surgery, trading signal precision for a far lower surgical burden.

ECoG: Places an electrode grid directly on the brain's surface, capturing cleaner signals than EEG without penetrating the cortex. It has enabled [speech decoding at 78 words per minute](#).

sEEG: Inserts thin depth electrodes into cortical and subcortical regions. Widely used to localize seizures, it has produced large clinical archives of human intracranial recordings.

Intracortical: Inserts microelectrodes into the cortex to record individual neurons, providing the highest-fidelity signals. Used by [Neuralink](#) and Utah arrays, it enabled Stanford's [90-character-per-minute handwriting](#)

[decoder](#) ▶ but surgery and long-term durability remain major barriers to scale.

The Future

NV-Center: Defects in diamonds that detect tiny magnetic fields at room temperature. Not yet practical for BCI, but huge potential if we make it work.

Appendix E: Glossary

ALS: Amyotrophic lateral sclerosis. Progressive motor-neuron disease; one of the conditions BCI work has historically focused on.

Bandwidth: How much information you can send and receive per unit of time.

Bandwidth gap: The vast mismatch between how fast machines can think and communicate, and how slowly humans can process information.

BIDS: Brain Imaging Data Structure. A community standard for organizing neuroimaging data (MRI, EEG, MEG) in a uniform directory layout. Adoption is still uneven.

Connectome: The full wiring diagram of a nervous system: every neuron and every connection between them.

Diffusion models: Image-generation models that learn to reverse a noise process: start with pure noise, denoise step by step into a coherent image.

Dura mater: The tough outer membrane enveloping the brain and spinal cord. Semi-invasive devices sit beneath the skull but rest on top of this layer.

EEG: Electroencephalography. Reads voltages off the scalp. Cheap, ubiquitous, noisy.

Endovascular: Electrodes delivered through blood vessels rather than open craniotomy.

Exocortex: An external digital layer of your mind that extends your memory, attention, and mental computation beyond your biological brain while functioning as part of it.

FAIR: Meta’s Fundamental AI Research lab. The open-research arm of Meta AI.

fMRI: Functional magnetic resonance imaging. Room-sized scanner with millimeter spatial resolution; the highest-quality non-invasive signal we have, at the lowest portability.

fNIRS: Functional near-infrared spectroscopy. Measures blood oxygenation through the skull using infrared. Cheap, wearable, slow.

Glial signaling: Communication between glial cells, the non-neuronal majority of brain cells. Increasingly suspected to modulate cognition, not just support neurons.

Hemodynamic: Relating to blood flow. Sensors such as fNIRS read the blood-oxygenation shadow of neural activity through the skull.

Interpretability: The effort to understand what happens inside a neural network: which internal features and circuits produce its behavior.

Limbic system: A set of deep brain structures that drive emotion, motivation, and reward. It is the circuitry that advertising has always tried to reach from the outside.

MEG: Magnetoencephalography. It reads the tiny magnetic fields cast by neural activity. Millisecond timing, but classically needs a room-sized cryogenic scanner; OPMs are a new wearable path.

Microelectrode arrays: Tiny silicon chips with dozens to hundreds of hair-thin needles, surgically inserted into brain tissue to record the firing of individual neurons.

Microtubules: Tube-shaped protein structures inside neurons. Some theories (Penrose-Hameroff) propose they support quantum computation; mainstream neuroscience is still skeptical.

Motor cortex: The strip of cortex that plans and executes voluntary movement. Easiest brain region to decode from, because intentions there map cleanly to body actions.

NWB: Neurodata Without Borders. A standardized data format and ecosystem for neurophysiology recordings. The closest thing to a common schema across labs.

OPM: Optically-pumped magnetometer. A chip-scale magnetic-field sensor that works at room temperature; the path to wearable MEG.

Phase Lock Window: The brief interval between recognizing that AI will surpass us and the moment it does; the span in which humanity still has the agency to shape the terms of the relationship. Roughly now to the mid-2030s.

Quantum effects: Hypothesis that quantum-mechanical phenomena (coherence, entanglement) play a functional role in brain computation. Live but speculative, no consensus evidence.

Ransomware: Malicious software that encrypts a victim's data or device and demands payment to restore access.

RLHF: Reinforcement learning from human feedback. The training step where humans rank a model's outputs to steer it toward preferred behavior.

sEMG: Surface electromyography. Reads electrical activity of muscles through skin electrodes. Well-suited to subtle motor intent and silent speech.

Signal-to-noise: How strong the wanted signal is against background noise; the core obstacle to reading the brain through skull and skin.

Whole Brain Emulation: A hypothetical process of scanning, mapping, and simulating a biological brain on a computer so that the digital model functions identically to the original mind.

wpm: Words per minute. A standard measure of communication speed.

VERSION HISTORY

2026-10-02 v1.0 · First publish.