

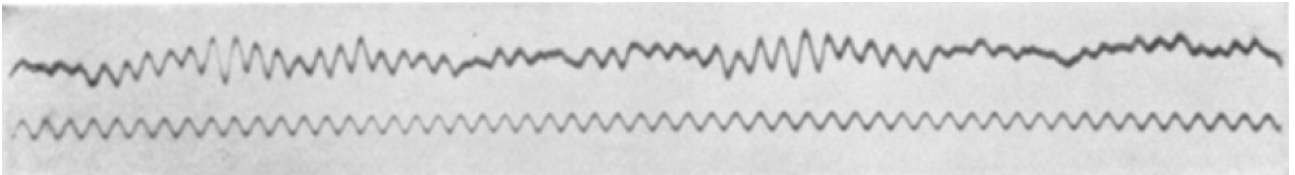
THE SETU VISION

# The Intent Layer

Machines learned to understand us faster than we learned to reach them. We think BCI is a few years from its ChatGPT moment. This is why, what it will take, and what Setu is building.

Short version: 2 minutes    Full version: 32 minutes    19 figures    62 sources

Written by the team at Setu



The first recording of electrical activity from a human brain, made by Hans Berger in 1924. Upper trace: the brain. Lower trace: a 10 Hz timing signal. A century later this is still the signal. Image: Hans Berger, public domain.

I

## The channel

Every computer still waits for a human body to type, tap or speak. That wait is now the slowest part of the system.

---

II

## The signals

Intent leaves physical traces from brain to muscle. Laboratories have already decoded handwriting, speech and images from them. EEG is the one sensor everyone could wear.

---

III

## The distance

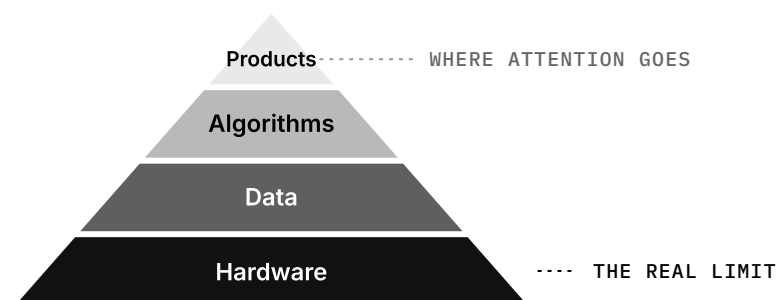
Decoding in a laboratory is not a product. The body is a superb incumbent, so the first wins come from silence, privacy and free hands, not from speed.

---

IV

## The learning problem

Hardware sets the ceiling, data is the fuel, and the winner is the system that learns each person fastest. That is why we bet on EEG, on scale, and on building the full stack.



---

**V**                      **The source: data and alignment**

Text, code and images are compressed products of thought. Neural data is the source. AI labs need it now, as training data and as a faster way to align their models. Those are the first two wedges.

---

**VI**                      **The road**

BCI reaches the world through two things people will actually wear: silent communication, and brain wellness and recovery. Its ChatGPT moment is a few years away.

---

**VII**                      **Agency**

None of this is worth building unless it can be switched off. Off by default, on the device, never for advertising, and the user always has the last word.

---

# The last mechanical step

Every computer, however capable, still waits for a human body to do something a sensor can register. That wait is now the slowest part of the system.

**S**omewhere tonight, a system that can draft a contract or re-plan a shipping route in seconds is doing nothing at all. It is waiting for someone to finish typing.

In a hospital ward, a woman who has had a stroke knows exactly what she needs to tell the nurse. The sentence is complete. It has no way out.

On a crowded train, a man has a message he would rather nobody overheard. He could say it in two seconds. He thumbs it out in forty.

An idle machine, a silenced patient, a self-conscious commuter. In each case the intelligence on either side is not the limit. The channel between them is.

**|** Computing has an input problem, and it is older than computing.

FIGURE 01 | SCHEMATIC

## One instruction, four ways across

The same intent can reach a machine by different physical routes. Each route asks something different of the body, and each stands on different evidence.

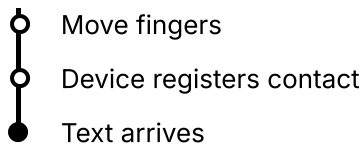
INTENT → **"Move Thursday's review to Friday"**

→ MACHINE ACTION

### Keys and touch

Mechanical input

🗨 Put it into words



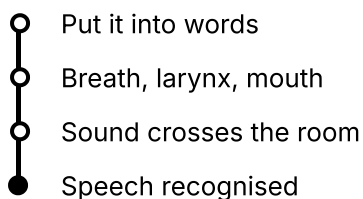
DEPLOYED EVERYWHERE

Hands and eyes occupied. Visible to anyone watching.

---

## Voice

Expressive input



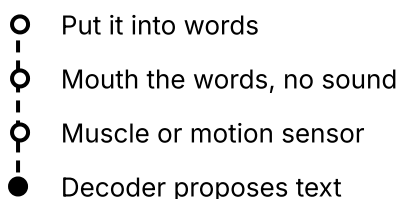
DEPLOYED WIDELY

Hands free. Audible to everyone nearby; needs a voice.

---

## Silent articulation

Biosignal input · peripheral, not brain



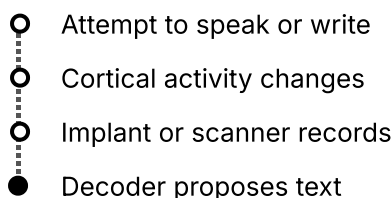
RESEARCH, EARLY DEVICES

Private and quiet. Still a deliberate movement; accuracy is vocabulary-dependent.

---

## Neural measurement

Neural input



CLINICAL TRIALS (IMPLANTS) WEARABLE FORM

Needs no working muscle. Best results today require surgery or a shielded room.

**How to read it.** Line style encodes maturity: solid = deployed at scale, dashed = demonstrated in research and early devices, dotted = demonstrated only under clinical or laboratory conditions. The moving pulse is decorative and encodes no speed data. **Sources:** Parts II–III. Schematic by the author.

# A century of asking the body for less

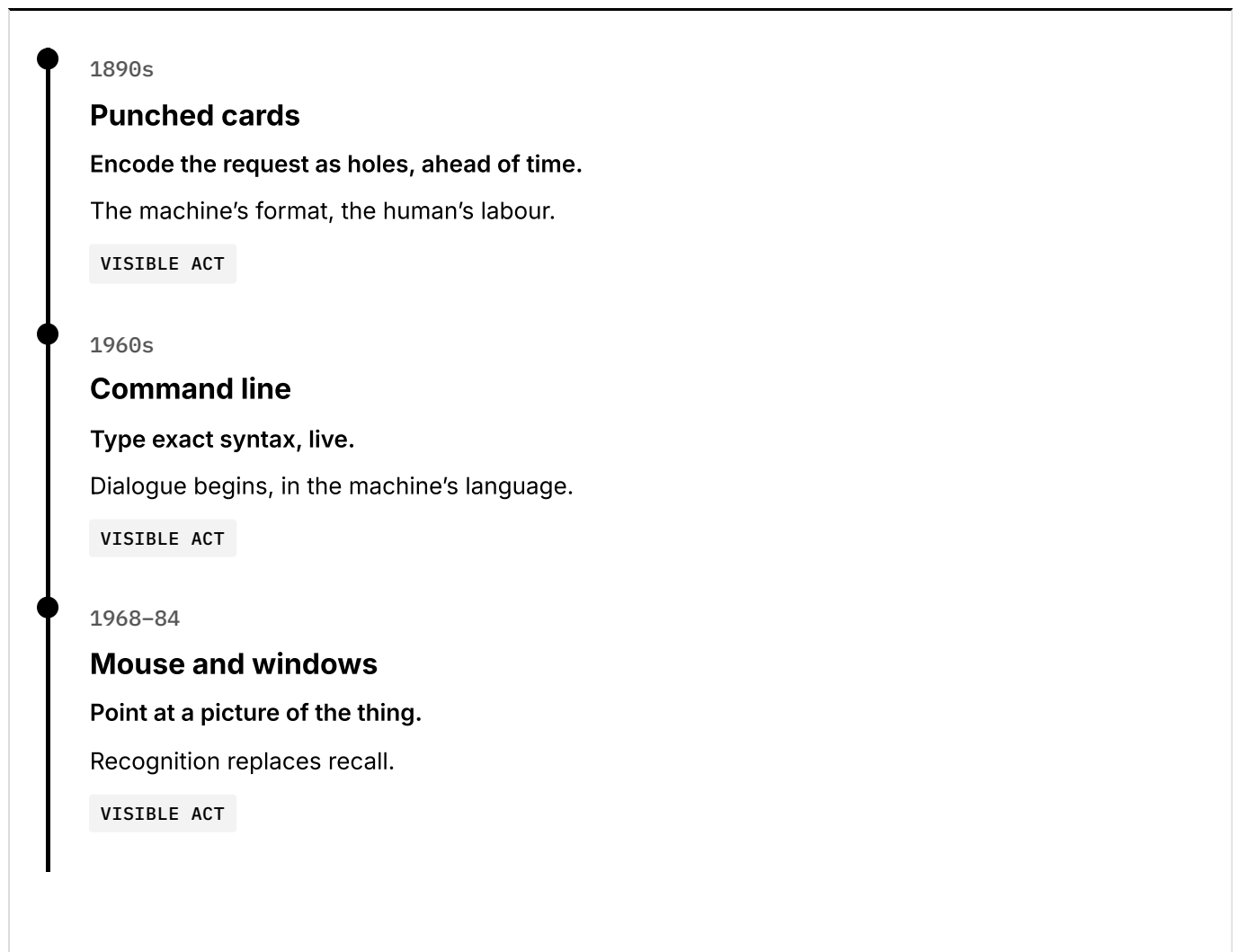
The history of the interface is a history of who does the translating. With punched cards, a person turned intent into holes, in advance, in the machine's format. The command line let them type that format live. Windows replaced syntax with pointing. Touch removed the pointer. Voice removed the screen. Language models removed the last of the machine's grammar: you can now say what you want in your own words and be understood.

Step by step, translation moved from the human to the machine. Transmission did not move at all. Every one of those interfaces still needs a hand that travels or a voice that carries.

## FIGURE 02 | TIMELINE

### What the human had to do

Each interface era cut the translation a person must perform. None, so far, has removed the need for an observable physical act.



2007

## Multi-touch

Touch the thing itself.

The computer goes everywhere the hand goes.

VISIBLE ACT

2011

## Voice assistants

Say a short command aloud.

Hands freed; privacy lost.

AUDIBLE ACT

2022

## Language models and agents

Describe the goal in your own words.

The machine finally learns our language.

VISIBLE OR AUDIBLE ACT

Proposed

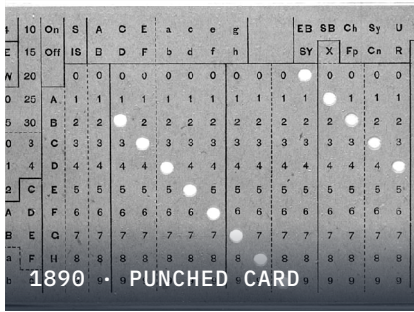
## Biosignal and neural input

Intend it deliberately; the body barely moves, or does not.

Not shipped at scale. Evidence in Part II.

UNOBSERVED ACT • PROPOSED

**Notes.** Dates mark when each mode became widely demonstrated or adopted (tabulating cards for census work, 1890; Engelbart's mouse demonstration, 1968, and mass-market graphical interfaces, 1984; capacitive multi-touch phones, 2007; mainstream phone voice assistants, 2011; public chat interfaces to large language models, 2022). The final column is a proposal, not a product category.



Three objects that changed what a person had to do to be understood by a machine. Photographs via Wikimedia Commons: Unknown author, public domain; The wub, CC BY-SA 4.0; The wub, CC BY-SA 4.0.

## The expression ladder

Sort inputs by what they ask of the body and four rungs appear.

1. **Mechanical.** You act on an object: a key, a mouse, a pane of glass.
2. **Expressive.** You act as you would toward a person: speech, gesture, gaze.
3. **Biosignal.** A sensor picks up the body preparing an act, such as a muscle firing, without the act being seen or heard.
4. **Neural.** An instrument measures the brain itself. This is a brain–computer interface, or BCI.

Each rung asks less of the body and is harder to make reliable. Rung three is not rung four. A sensor on the jaw reads muscle. It does not read mind.

Why climb? Because for some people the mechanical step is a wall: paralysis and stroke can leave language intact and its exits closed. Because speech is a broadcast, and much of what people want to tell a computer is not for the room. Because hands are often full. And because once software can act for us, the instructions and corrections we give it become the most valuable input in the system.

## The ChatGPT moment for BCI

Not long ago, most researchers avoided the phrase “general intelligence” in print. It was a taboo. It now appears in corporate mission statements, and the world was wrong.

It was wrong in a particular way. In 2017 a paper introduced the architecture every large language model now uses, and almost nobody outside the field noticed. For five years the capability compounded in plain sight. Then a text box put it in front of the public, and the consensus about what machines could do collapsed in a season.

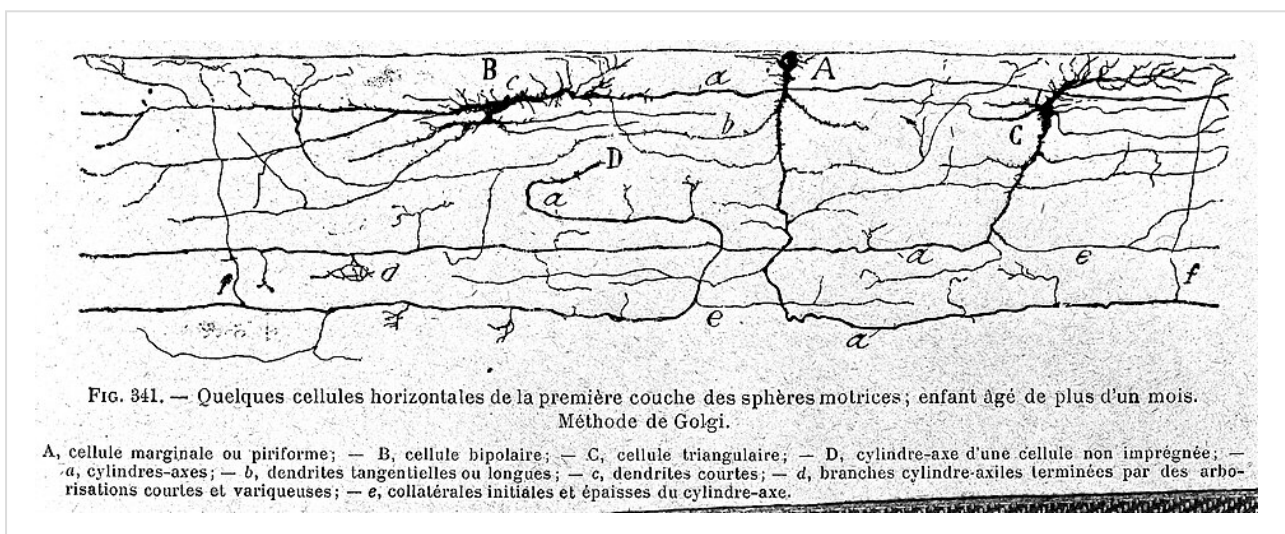
■ We think BCI today is where AI was in 2017.

Set the two histories side by side. In 2012 deep networks suddenly worked; in 2021 a paralysed man typed from his motor cortex. In 2017 an architecture arrived that could scale; today the first general models for brain data are being trained. In AI, the first large models and the first interface anyone could use followed within five years.

We expect the same order here, at a similar pace. BCI is a few years from its ChatGPT moment: the day a device with no surgery does something useful enough that millions of people put it on. Most people are underestimating that day for the same reasons they underestimated AI.

# What the body already broadcasts

Intent leaves physical traces all the way from cortex to air. The question is which trace you can catch, and where you have to stand to catch it.



The cerebral cortex, drawn cell by cell from a stained section. Every method in this part listens, from a different distance, to tissue like this. Image: Wellcome Collection, CC BY 4.0.

When you decide to say a word, neurons in the cortex change their firing. Currents flow. Seconds later, blood supply shifts to feed the active tissue. Commands descend to the muscles of the tongue, jaw and larynx. Tissue moves. Air vibrates. Each stage is a place where an instrument can listen.

A BCI measures brain activity and translates it into commands for a device, giving the brain an output that does not depend on muscle.<sup>4</sup> The term was coined in 1973.<sup>5</sup> This is an old science with new tools.

Three rules make any result in this field easier to read. Reading is not writing: recording the brain and putting information into it are different problems, and this essay is about reading. Movement is easier than meaning: the strongest results decode attempted

handwriting and speech. And the verb matters: seeing, hearing, attempting and imagining are different brain states, so always ask what the participant was doing.

## Nine ways to listen

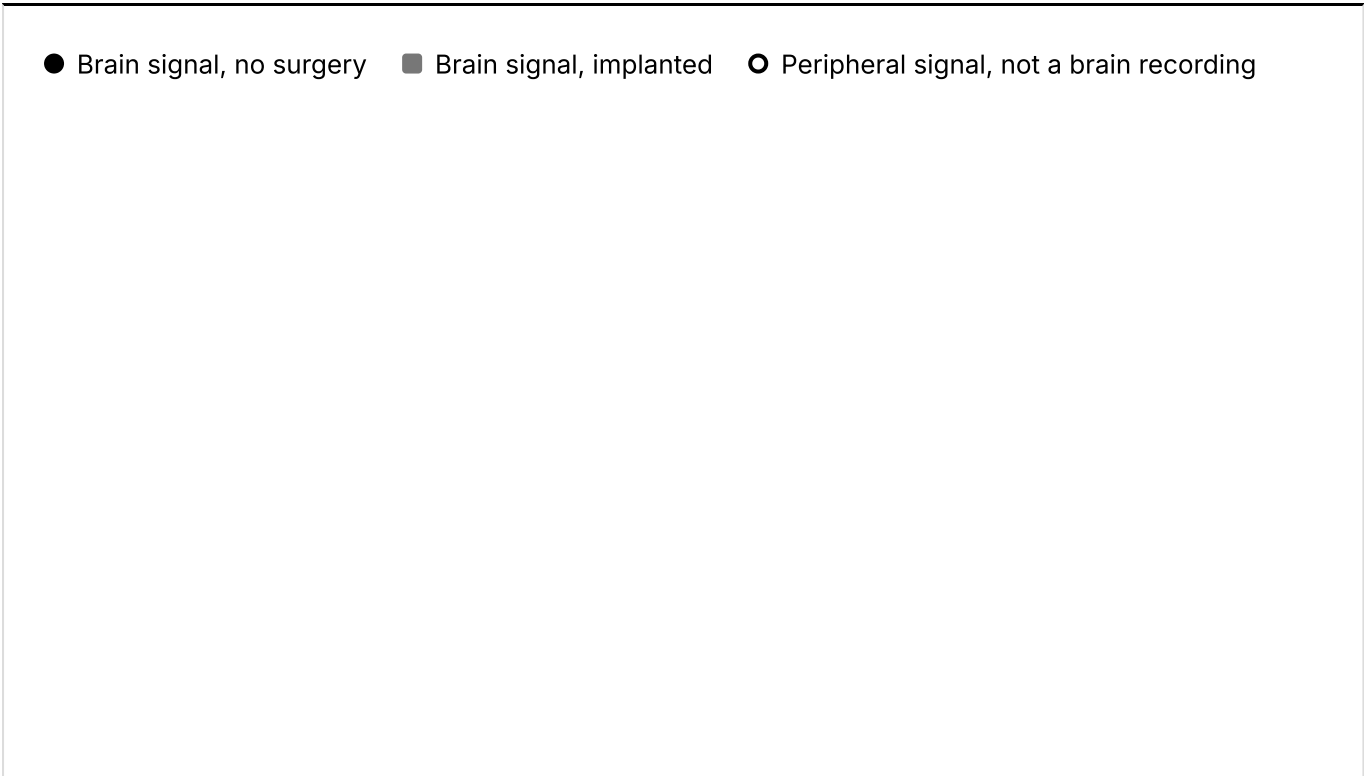
A scalp electrode stands to a neuron roughly as a weather satellite stands to a raindrop. It cannot see the single event. It sees the organised behaviour of millions, blurred by the skull. An implanted electrode is a rain gauge: exact, local, and somebody has to go through bone to install it.

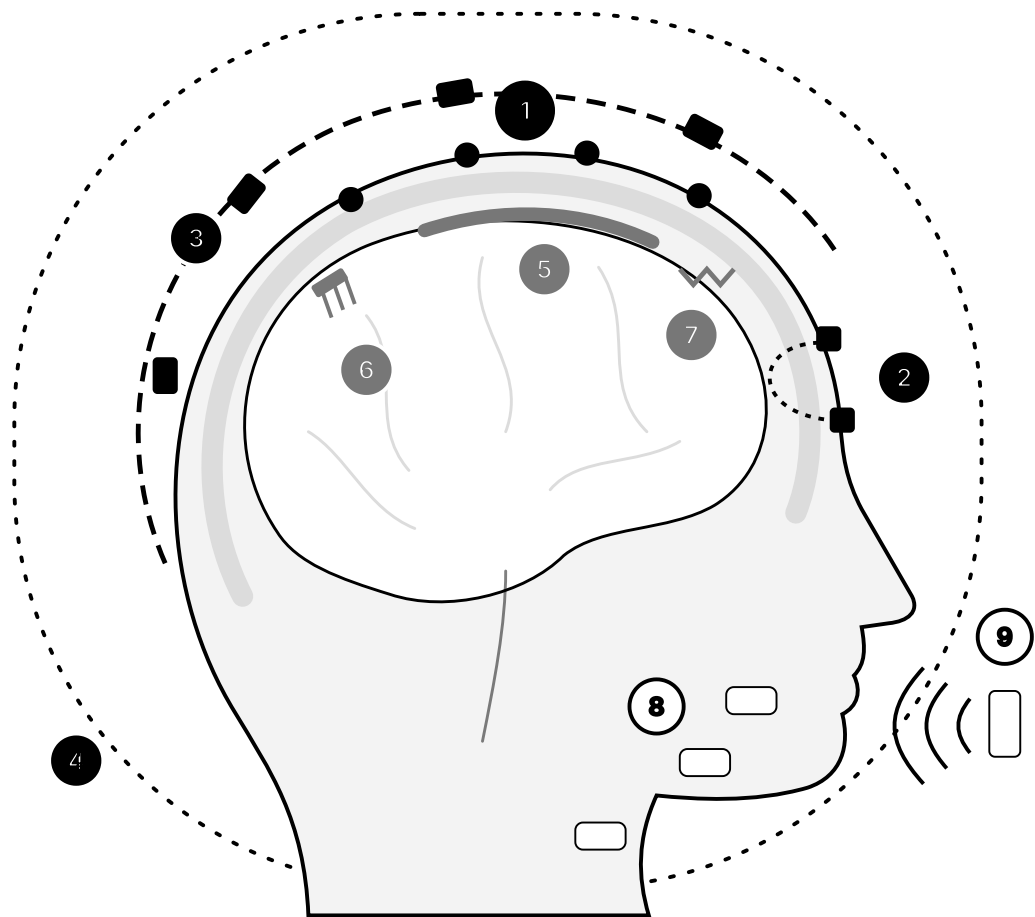
Satellites changed the world and rain gauges did not, for one reason. Satellites cover everyone. EEG, the recording of electrical activity at the scalp, is the satellite of this field. It is a century old, safe, light and cheap enough to build anywhere. That is why it is the instrument this essay keeps returning to.

FIGURE 03 | INTERACTIVE

### Where each instrument listens

Seven methods record the brain from different depths; two record the periphery and never touch neural tissue. Select a number to see what each does and does not measure.

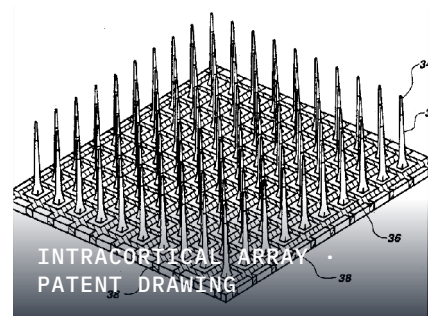
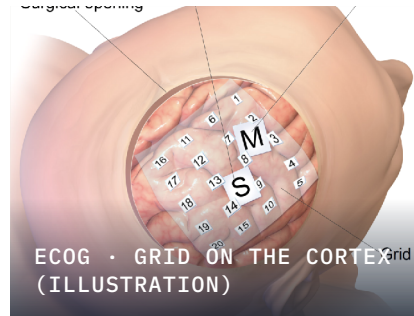
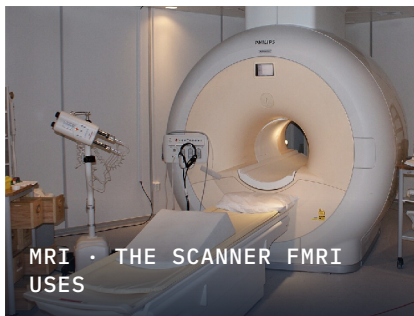
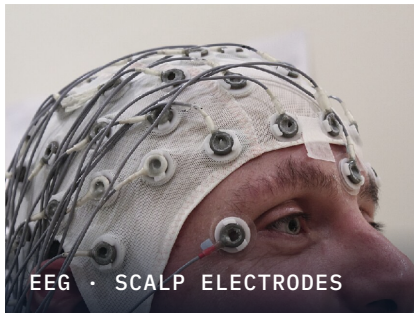




## 1 • Scalp EEG BRAIN • NO SURGERY

|           |                                                                                                                               |
|-----------|-------------------------------------------------------------------------------------------------------------------------------|
| MEASURES  | Voltage at the scalp from large populations of cortical neurons firing in step. Millisecond timing.                           |
| DOES NOT  | Resolve small or deep sources. Skull and tissue smear the signal; muscle and eye movement easily swamp it.                    |
| PRACTICAL | Cheap, portable, decades of clinical use. Dry electrodes trade convenience for signal quality. Hair, sweat and motion matter. |

**Schematic, not anatomical scale.** Sensor positions are illustrative. The dotted outline around the head stands for an fMRI scanner bore; the dashed arc for an MEG sensor helmet. Descriptions summarise standard measurement physics; study citations are in Figure 4.



The instruments themselves. The first four involve no surgery; the last two are implanted. Images via Wikimedia Commons: Chris Hope, CC BY 2.0; Alexarje, CC BY 4.0; National Institute of Mental Health, public domain; Jan Ainali, CC BY 3.0; Blausen Medical (BruceBlaus), CC BY 3.0; Richard A. Normann, patent drawing, public domain.

## The evidence, verb by verb

**Attempting to write.** In 2021 a man paralysed from the neck down imagined writing letters by hand while implanted arrays recorded his motor cortex. The system produced 90 characters a minute at 94.1% accuracy.<sup>6</sup>

**Attempting to speak.** In 2023, two implant studies decoded attempted speech at 78 and 62 words a minute, with about one word in four wrong.<sup>7 8</sup> A year later a third held 97.5% accuracy across eight months of daily conversation.<sup>9</sup>

**Attempting to move.** People with paralysis have steered robotic arms,<sup>20</sup> felt touch written back into the brain,<sup>21</sup> and walked again through a bridge between cortex and spinal cord.<sup>23</sup> EEG alone has driven a robotic arm in continuous tracking.<sup>22</sup>

**Perceiving.** From inside a scanner, models have recovered the gist of a story a person was hearing,<sup>11</sup> rebuilt pictures they were looking at,<sup>12</sup> and adapted to a new person with one hour of data.<sup>13</sup> With EEG and MEG, models can pick out which stretch of speech a listener was hearing.<sup>17</sup> These pair a recording with a generative model. The signal supplies the meaning; the model paints the detail.

**Typing.** With no surgery, sentences have been decoded from magnetic recordings of people typing, at 39% word error.<sup>16</sup> The same task on EEG is harder today: two characters in three were wrong.<sup>15</sup> We read that as the starting line. Every curve in Part IV says EEG has the most to gain from data.

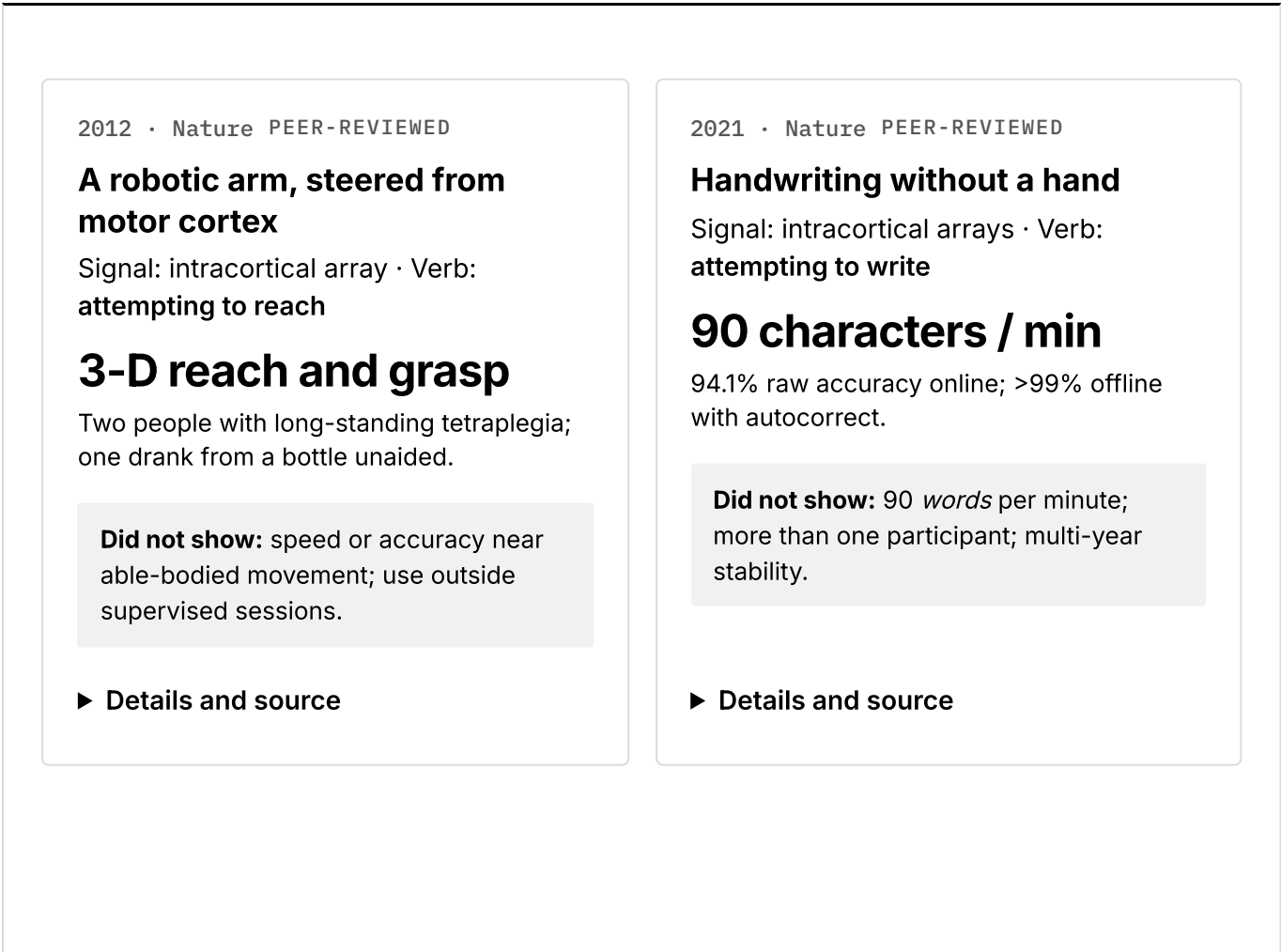
**Mouthing.** Away from the brain, silently mouthed words have been turned into speech from facial muscle signals,<sup>18</sup> and a wrist sensor trained on thousands of people now works on strangers with no calibration.<sup>19</sup>

**Thinking in words.** In 2025, imagined sentences were decoded from four implanted participants. The same study showed how to stop a device decoding inner speech the user had not offered.<sup>25</sup>

FIGURE 04 | EVIDENCE

## The evidence wall

Sixteen studies, each stated the same way: what was measured, what the person was doing, what came out, and what the result does not show.



2021 · Science PEER-REVIEWED

## Touch, written back to the brain

Signal: intracortical recording + stimulation · Verb: **attempting to grasp**

### 20.9 s → 10.2 s

Median task time with a robotic arm, without and with evoked tactile sensation.

**Did not show:** rich or naturalistic sensation; writing of language or concepts.

► [Details and source](#)

2023 · Nature PEER-REVIEWED

## Speech, voice and a face from the cortical surface

Signal: ECoG grid · Verb: **attempting to speak**

### Median 78 words / min

Median 25% word error rate; text, synthetic voice and avatar.

**Did not show:** error rates low enough for unassisted daily use; decoding of unattempted thought.

► [Details and source](#)

2023 · Nature PEER-REVIEWED

## Large-vocabulary speech from single neurons

Signal: intracortical arrays · Verb: **attempting to speak**

### 62 words / min

23.8% word error on a 125,000-word vocabulary; 9.1% on 50 words.

**Did not show:** conversational accuracy (see 2024); results beyond one participant.

► [Details and source](#)

2024 · NEJM PEER-REVIEWED

## A speech prosthesis used for months

Signal: intracortical arrays · Verb: **attempting to speak**

### 97.5% word accuracy

Sustained over 8.4 months; ~32 words / min; 248+ hours of self-paced use; 30-minute first calibration.

**Did not show:** generalisation across patients; performance without ongoing recalibration.

► [Details and source](#)

2025 · Cell PEER-REVIEWED

## Inner speech, and a lock for it

Signal: intracortical arrays · Verb: imagining speech

## Imagined sentences, real time

Inner speech is represented in motor cortex; some unprompted inner speech was decodable; safeguards blocked it.

**Did not show:** decoding of arbitrary private thought; anything outside implanted motor cortex.

► Details and source

2023 · Nat. Neurosci. PEER-REVIEWED

## The gist of a story, from blood flow

Signal: fMRI · Verb: listening (also imagining, watching)

## Meaning, not words

Generated word sequences that recover the sense of heard speech.

**Did not show:** word-for-word transcripts; decoding without the person's cooperation, which was required to train and to apply the decoder.

► Details and source

2023 · NeurIPS PEER-REVIEWED

## Rebuilding a viewed image (MindEye)

Signal: 7-tesla fMRI · Verb: looking at pictures

## Retrieval and reconstruction

Identifies the viewed image among similar candidates; generates a resembling image with a diffusion model.

**Did not show:** imagined or remembered images; exact readout. The generator supplies detail the recording lacks.

► Details and source

2024 · ICML PEER-REVIEWED

## A new brain with one hour of data (MindEye2)

Signal: 7-tesla fMRI · Verb: looking at pictures

## ~1 hour to adapt

After pretraining across 7 other people in the same dataset.

**Did not show:** zero-calibration transfer; transfer across scanners, tasks or to portable sensors.

► Details and source

2024 · NeurIPS PEER-REVIEWED

## Short video from a scanner (NeuroClips)

Signal: fMRI · Verb: watching video

## Clips up to 6 s at 8 fps

Resembling the watched footage, via a pretrained text-to-video model.

**Did not show:** a live feed, imagined scenes, or frame-accurate recovery.

► Details and source

2025 · arXiv PREPRINT

## Typed sentences from outside the skull (Brain2Qwerty)

Signal: MEG and EEG · Verb: typing  
memorised sentences

## 32% vs 67% character error

Average for MEG vs EEG. Best MEG  
participants: 19%.

**Did not show:** free-form thought to text.  
Participants were physically typing; EEG  
got two characters in three wrong.

► Details and source

2026 · arXiv PREPRINT

## More data, fewer errors (Brain2Qwerty v2)

Signal: MEG, real time · Verb: typing  
natural sentences

## 39% word error rate

Average; improves log-linearly with data.  
Best participant: half of sentences with at  
most one wrong word.

**Did not show:** anything wearable (MEG  
needs shielding); peer review is  
pending; ~10 hours of recording per  
person.

► Details and source

2025 · arXiv PREPRINT

## The same task on four machines

Signal: EEG, MEG, 3T and 7T fMRI · Verb:  
looking at pictures

## 498 hours, 84 people

The most precise devices decode best. No  
plateau with more data. Hours per person  
matter more than more people.

**Did not show:** where any device's  
ceiling lies; anything about language or  
intent.

► Details and source

2020 · EMNLP PEER-REVIEWED

## A voice for silently mouthed words

Signal: facial surface EMG · Verb: mouthing words silently

**64% → 4% word error**

Closed-vocabulary condition. Open vocabulary: 88% → 68%.

**Did not show:** brain decoding of any kind; usable open-vocabulary accuracy; multi-speaker generalisation.

► [Details and source](#)

2025 · Nature PEER-REVIEWED

## A wristband that works on strangers

Signal: wrist surface EMG · Verb: gesturing and handwriting

**20.9 words / min**

Handwriting, with no per-user calibration; +16% with personalisation.

**Did not show:** speech or brain decoding. Requires functioning hand muscles.

► [Details and source](#)

**Reading rules.** Numbers are quoted from each paper's abstract in its own metric; they are not comparable across cards. Preprints have not completed peer review. No stimulus or reconstruction images are reproduced here because image rights rest with the original authors and datasets; follow the source links to see them, including failure cases.

The laboratory has proved that the information exists. It has not yet proved that it travels.

# Decodable is not wearable

The incumbent is not the keyboard. It is the human body, which is fast, reliable, free and already installed.

A new interface competes with the hand, the voice and the glance. They need no charging, no calibration and no terms of service. The average person types about 50 words a minute<sup>1</sup> and speaks about 160.<sup>8</sup> Across seventeen languages, fast talkers and slow ones converge on the same rate of information.<sup>2</sup> No language is the default, and an interface tuned to one is unfinished.

## Four gates

Between a signal and a product stand four tests, passed in order.

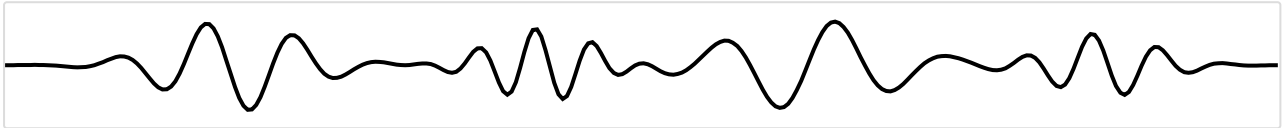
1. **Information.** Is the content physically present in what the sensor measures?
2. **Reliability.** Does decoding survive movement, sweat, hair, fatigue and a sensor that sits two millimetres to the left of yesterday?
3. **Utility.** Do accuracy, delay, setup, comfort and price add up to something better than the alternative at hand?
4. **Agency.** Can the person switch it on and off, see what was inferred, and reject it?

Part II was a tour of the first gate. Products are made or lost at the other three.

## The journey of a signal

Decoding is a chain, and the weakest link is usually the first. Degrade the signal and watch what an honest decoder should do, then let a language model paper over the gap.

PATTERN THE BODY PRODUCED • ILLUSTRATIVE



WHAT THE SENSOR RECORDED



AFTER ARTIFACT REMOVAL AND FILTERING



DECODER OUTPUT

**Move Thursday's review to Friday**

SUPPORTED BY SIGNAL

The cleaned trace still resembles the source. The decoder proposes one reading; the user confirms.

- 01 Sensor and synchronisation
- 02 Artifact removal
- 03 Learned representation
- 04 Task decoder
- 05 Language prior
- 06 Uncertainty estimate
- 07 Personal calibration
- 08 Deliberate intent gate
- 09 Editable output and feedback

**Synthetic demonstration.** All traces are generated in your browser from sine bursts and pseudo-random noise. They are not recordings and the slider is not calibrated to any device or measured signal-to-noise ratio.

The three output states are scripted to illustrate a principle: when the signal no longer supports a reading, a fluent sentence produced by the language model is *supplied*, not *measured*.

## Two clocks

A software agent can read a thousand pages while you read this sentence. Then it stops and asks a question, and the clock that matters changes. Its clock runs in milliseconds. Yours runs in the time it takes to understand, decide, find words and get them out.

The pressure this creates is easy to predict. When the human turn is the slowest step, there is a standing incentive to remove it. People are rarely pushed out of the loop by decree. They are pushed out by impatience: the approval that becomes a rubber stamp, the review that is skipped because the queue is long.

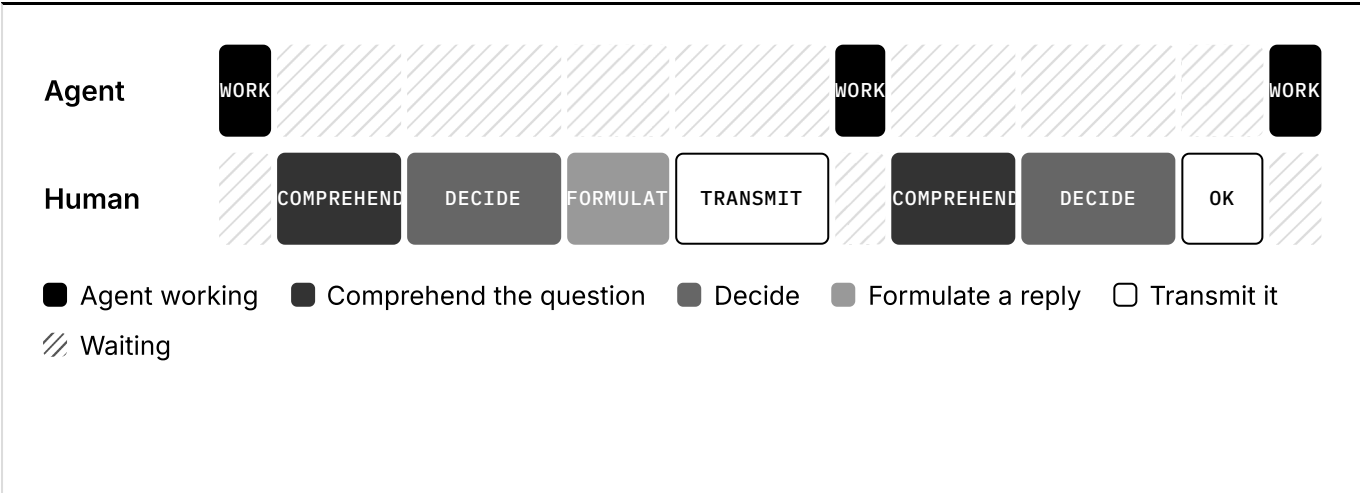
Whatever makes a human turn cheaper makes it likelier to survive.

Better prompts, better dashboards and more autonomous agents all work around the bottleneck. Only an interface that picks up intent closer to its origin attacks it. Most of what is proposed for keeping AI in check works by constraining the machine. BCI is one of the very few levers that works the other way, by raising what the human can do.

FIGURE 06 | INTERACTIVE

### Two clocks, one task

An agent finishes its part quickly and waits. The human turn is four activities, and a faster input channel shrinks only one of them.



With today's channel, the agent is idle for most of the task. Transmission is a visible share of the human turn, but not the largest.

**Illustrative proportions, not measurements.** Segment widths are chosen to show structure and are not drawn from data; real ratios vary enormously by task and person. The point is structural: comprehension and decision remain when transmission shrinks. The one published throughput estimate cited in the text<sup>3</sup> is not plotted because it does not decompose into these stages.

## Different thresholds, different lives

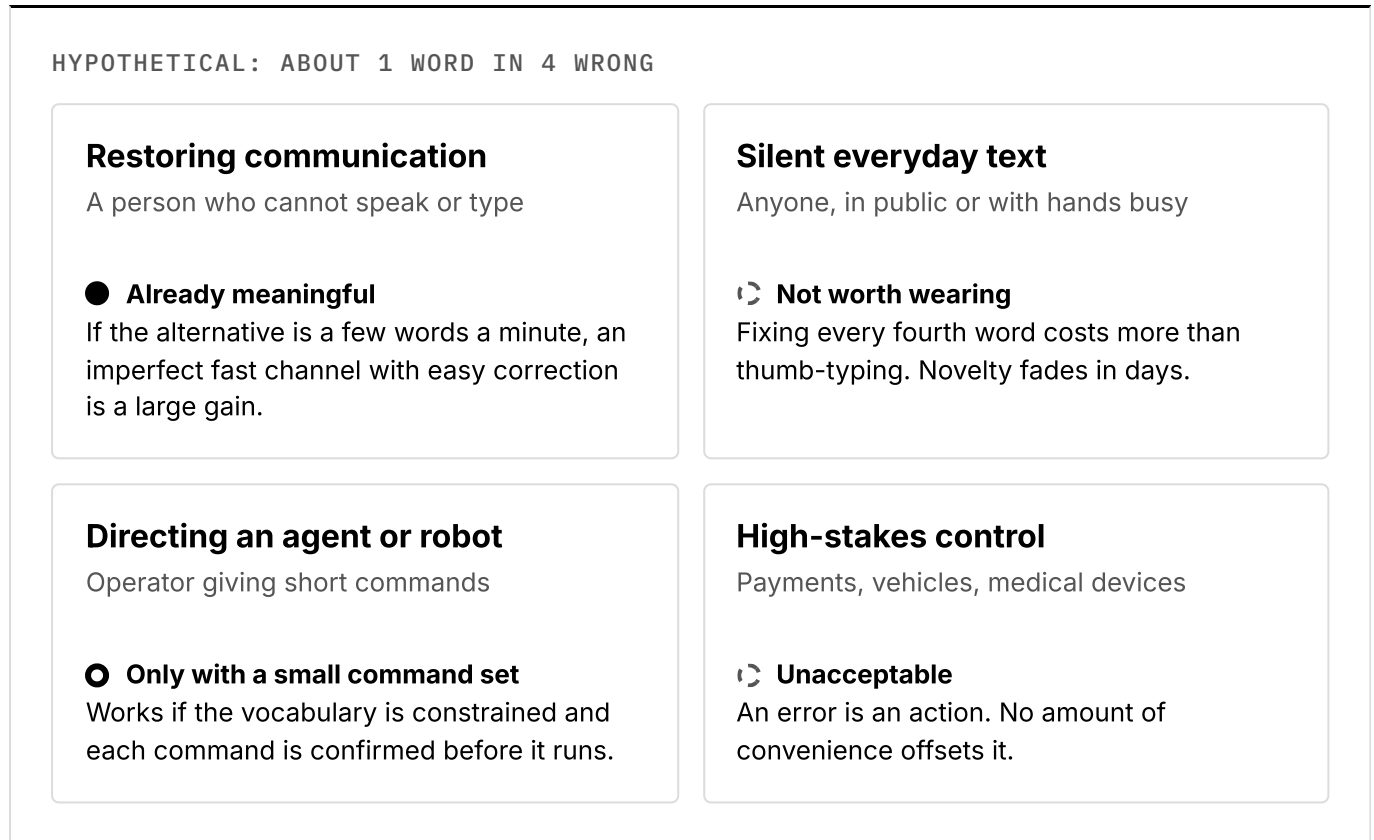
For someone who communicates through an eye-tracker at a few words a minute, thirty words a minute with errors changes what a day contains. The people in the trials above accepted brain surgery for that chance. They are who this science exists for, and they are owed finished technology, not prototypes.

For someone with working hands, thirty words a minute is a downgrade. A consumer interface will not win on speed first. It will win on silence, privacy, free hands and access.

And the bar is a cliff, not a slope. Speech recognition that is right nineteen times in twenty is a toy. Right nine hundred and ninety-nine times in a thousand, it is everywhere. A field can look stalled for years and then tip.

## The same error rate, four verdicts

Pick a hypothetical decoder. Its value changes completely with who is using it and what a mistake costs.



**Author's judgment, not data.** The three error levels are hypothetical and are not claims about any device or about "where the field is". Verdicts are qualitative assessments that assume a workable correction mechanism; they are offered to show that usefulness is a surface over need, burden and consequence, with no single threshold.

# The interface is a learning problem

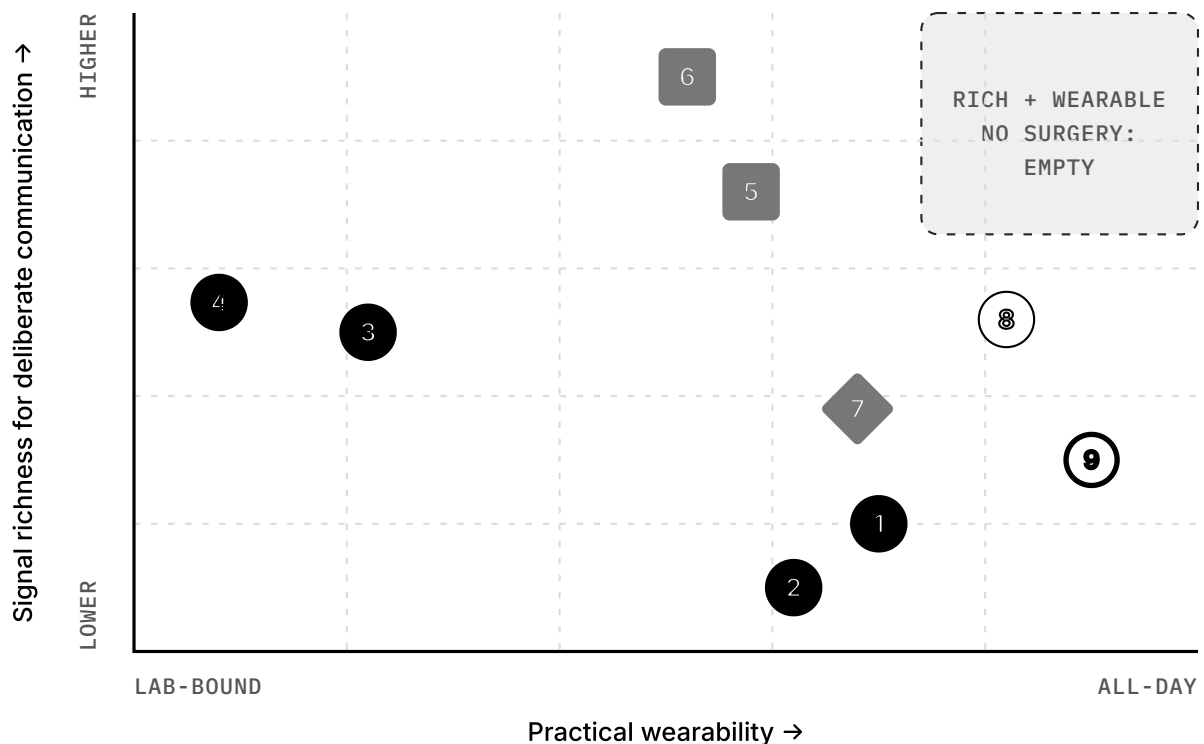
Hardware is the foundation, data is the fuel, and the system that wins is the one that learns each person fastest. Only EEG can do all three at the scale of a population.

Implants give the richest signals and every headline result for speech. Each one is also a neurosurgical operation. Their reach is bounded by the number of surgeons and by how much risk a person's situation justifies. For severe paralysis that is clear. For everyone else it is not, and it will not be for a long time.

A device without surgery can be redesigned in weeks, tested on thousands of people and shipped without a hospital. So the real question is narrow, and experiments can answer it: how much useful information can be captured from outside the skull, and how cheaply can a model be adapted to a new person on a new day?

## The unoccupied corner

Plot every sensing method by how wearable it is and how rich a signal it yields for deliberate communication. One corner of the map is empty. We think EEG, trained at scale, is what moves into it.



● Brain, no surgery   ■ Open surgery   ◆ Catheter-delivered   ○ Peripheral, not brain

- 1 **EEG.** Wearable; richness limited by the skull. Robust for a few commands, weak for open text so far.
- 2 **fNIRS.** Wearable; slow blood signal carries little per second.
- 3 **MEG.** Best non-invasive language results, inside a shielded room.
- 4 **fMRI.** Rich in space, slow in time, immobile.
- 5 **ECoG.** Goes everywhere once implanted; getting there is brain surgery.
- 6 **Intracortical.** Richest signal recorded; surgical, small coverage, durability open.
- 7 **Endovascular.** Less surgery, fewer channels, small command sets to date.
- 8 **sEMG.** Rich for deliberate movement only; needs working muscle. Not a brain signal.
- 9 **Radar / motion.** Contactless; early evidence; senses movement, not intent.

**Ordinal positions, author's judgment.** Neither axis has units. Placement reflects the published demonstrations in Figure 4 and standard measurement physics, not a measured benchmark; reasonable experts would move several points. "Wearability" for implants describes daily use after surgery; marker shape carries the surgical burden. Peripheral sensors are plotted for comparison and are a different kind of signal, not a point further along the same road.

## The ceiling is set at the sensor

Picture the field as four layers. Hardware at the bottom: what the sensor registers. Then data: what has been recorded. Then algorithms. Then products. Attention and money pool at the top, where the demonstrations live. The constraint runs the other way.

Information theory states the reason bluntly. No computation on a signal can add information about its source. A decoder can waste what the sensor captured. It cannot create what the sensor missed.

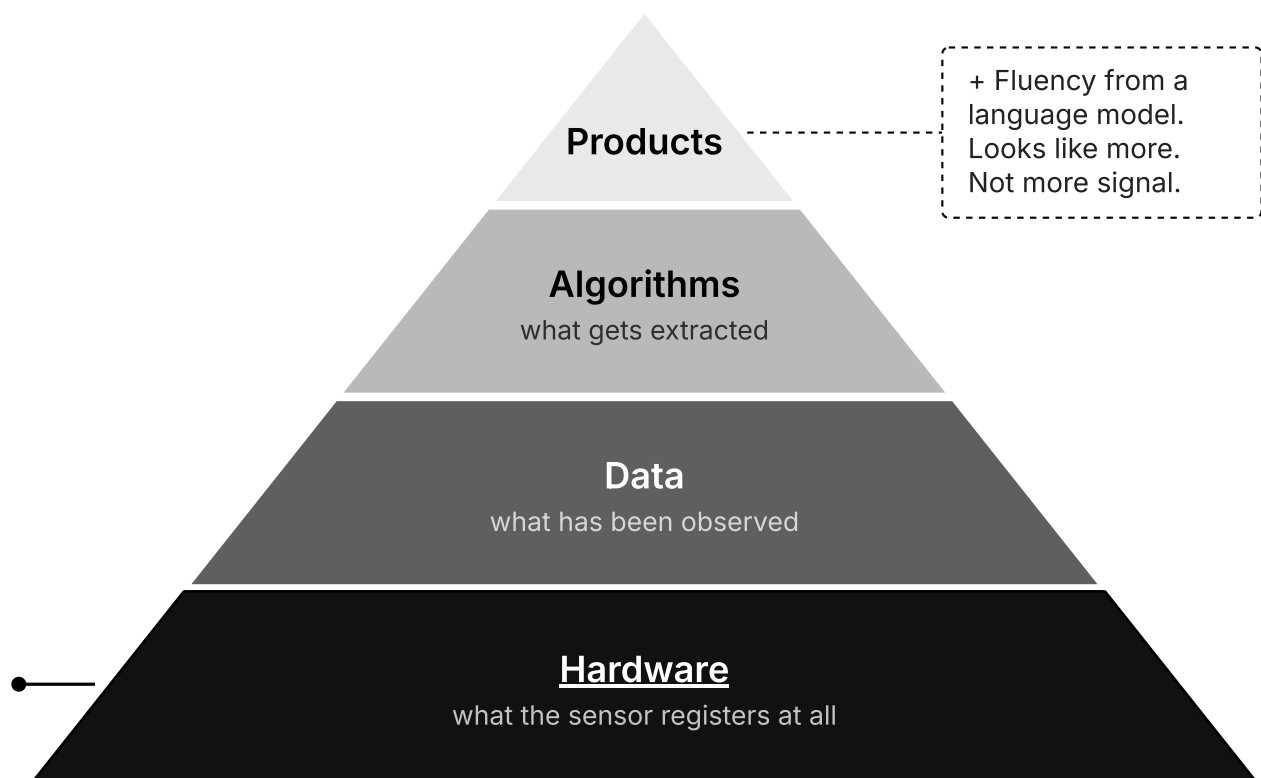
Hardware is the actual bottleneck, and product is what people are focusing on. That mismatch is the state of the field.

FIGURE 09 | INTERACTIVE

### Four ceilings

Each layer limits the one above it. Information about what a person intended can be lost on the way up and never gained. Select a layer.

↑ INFORMATION ABOUT INTENT CAN ONLY SHRINK



## Hardware sets what exists

|                    |                                                                                                        |
|--------------------|--------------------------------------------------------------------------------------------------------|
| BOUNDS             | The information present in the recording. Everything above works with this and nothing else.           |
| CANNOT BE FIXED BY | More data or a larger model.                                                                           |
| EVIDENCE           | On one task across four devices, the more precise sensor decoded better at matched data. <sup>34</sup> |

**Conceptual.** Layer widths are illustrative, not measured. The ordering rests on a theorem, the data-processing inequality: no processing of a signal can increase the information it carries about its source. The hatched block marks what a language prior contributes, which is plausibility rather than measurement.

## Why EEG wins

One experiment ran the same task, working out which image a person is looking at, on four kinds of machine: EEG, MEG and two strengths of fMRI. Eight datasets, 84 people, 498 hours.<sup>34</sup> Two results matter. Deep learning helped most on the noisiest sensor, which is EEG. And no sensor showed a plateau as data grew.

MEG has a cleaner view and needs a magnetically shielded room.<sup>30</sup> A technology that cannot leave a special facility cannot learn from a million people. fNIRS is wearable, but it measures a blood response that arrives seconds late, too slow for anything at the speed of speech.

EEG is fast enough, cheap enough and safe enough to be worn by anyone, anywhere. Its weakness is noise, and noise is exactly what modern learning is best at overcoming. At equal data the precise instruments lead. But the data will never be equal. A sensor that costs a thousand times less and goes home with the user gathers thousands of times more, and the history of machine learning is the history of the scalable method beating the elegant one.

Building EEG for scale changes the design. Portability comes first. Other sensors that do not interfere are stacked beside it. Sensors and models are designed as a pair, so the model's results decide what the next sensor measures. Cameras and microphones sit on the device, because context is what gives a recording its meaning.

## A field without its ImageNet

Modern AI was built on collections nobody had to commission. The image dataset that launched deep learning gathered millions of photographs already on the web.<sup>59</sup> A speech recogniser released in 2022 learned from 680,000 hours of audio collected from the internet.<sup>42</sup>

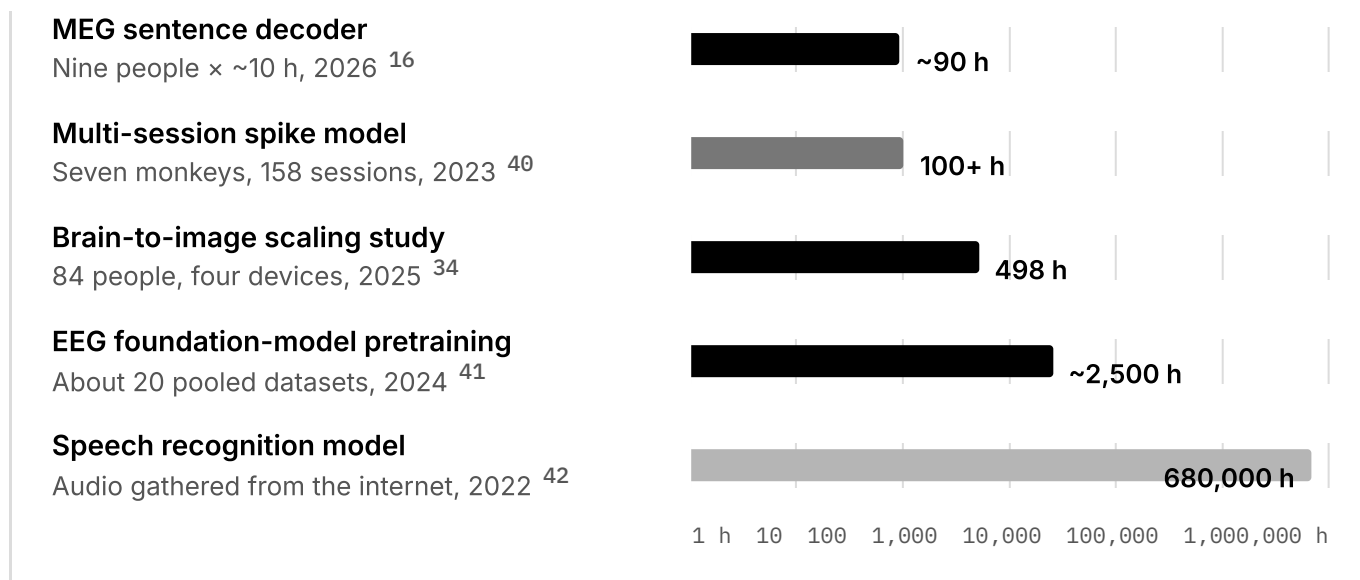
Brain data cannot be scraped. Every hour needs a consenting person under a sensor. So this field counts in hours where its neighbours count in lifetimes.

FIGURE 10 | DATA

### Hours, not lifetimes

The largest neural datasets behind published decoders are hundreds to thousands of hours. A single speech model was trained on 680,000.





**Logarithmic scale: each gridline is ten times the last.** Hours of recording used for training, as stated in each paper's abstract; the fMRI language figure is per participant and cited from the published record. Black = brain recording without surgery, grey = implanted, light grey = audio. An hour of audio and an hour of brain recording do not carry the same information; the comparison is of collection scale only.

Public datasets exist and matter.<sup>29 57 58</sup> They differ in sensor, layout, task and file format. Pooled, they are a patchwork. The field grew up inside medicine, where the habit is a small careful study and an algorithm built by hand for one patient.

There is also a trap. Hardware makers will not scale until there is demand for data. Data collectors cannot gather volume until hardware is cheap. Model builders cannot justify large training runs until the data exists. Each layer is sensibly waiting for the one beneath it.

The way out is to build every layer at once, ahead of demand. That is the case for a full-stack hardware company, and it is the company Setu is building. The one biosignal system that already works on strangers came from exactly this kind of integrated effort.<sup>19</sup> It does not have to be a solitary one: shared formats and shared benchmarks make every stack better.

## Scale is the strategy

AI learned this the hard way, and one of its founders named it the bitter lesson: general methods that scale with data and computation end up beating methods built on human cleverness.<sup>61</sup> BCI does not need seventy years to learn the same thing.

The measurements agree. Decoding improves steadily as recordings grow, for images<sup>34</sup> and for sentences.<sup>16</sup> Larger EEG models do not yet beat small ones, which tells us the models are starved, not that they are wrong.<sup>26 27 28</sup> And more hours from one person help more than more people do,<sup>34</sup> which is an argument for a device worn every day.

None of this requires understanding the brain first. Nobody fully knows how it computes. The wager is that enough structure can be pulled from its signals without a complete theory, as it was for language. Nor does it require new mathematics. The methods were worked out once, for language models, and they carry over.

FIGURE 11 | TABLE

### The playbook, transplanted

Ten techniques that turned pretrained AI models into useful ones, each with its counterpart in biosignal decoding, the state of the evidence, and what goes wrong.

| IN AI                              | COUNTERPART IN BIOSIGNAL DECODING                                                                                                                 | THE CATCH                                                                                               |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| <b>Self-supervised pretraining</b> | Train on unlabelled recordings pooled across people and sessions: masked-signal models for EEG, multi-session models of spiking. <sup>41 40</sup> | In benchmarks, larger models were not reliably better and specialists stayed competitive. <sup>27</sup> |
| <b>Supervised fine-tuning</b>      | Train on a labelled task with known ground truth: cued speech, typed sentences, viewed images.                                                    | Labels need a ground truth. For private mental states there is none to check against.                   |
| <b>Transfer across people</b>      | A shared model plus a light alignment per person; <sup>13</sup> a generic wrist-signal model. <sup>19</sup>                                       | For image decoding, adding people helped little; adding hours per person helped. <sup>34</sup>          |
| <b>Few-shot personalisation</b>    | Minutes to an hour of a new person's data. <sup>9 13</sup>                                                                                        | The burden lands on the user, and returns if drift is unmanaged.                                        |
| <b>Continual learning</b>          | The decoder keeps updating during ordinary use; one ran 403 days without a retraining session. <sup>35 9</sup>                                    | Self-training on corrected guesses can drift toward the language model.                                 |

| IN AI                                         | COUNTERPART IN BIOSIGNAL DECODING                                                                                     | THE CATCH                                                                         |
|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| <b>Test-time adaptation</b>                   | Re-align each day's activity to a stable underlying structure, with no labels. <sup>36 37</sup>                       | Assumes the structure is stable. Unproven for signals recorded through the skull. |
| <b>Language prior and search at inference</b> | A language model rescoring the decoder's candidates. <sup>8 9</sup>                                                   | Adds fluency, not information. Report results with and without it.                |
| <b>Learning from human feedback</b>           | Corrections and confirmations as labels; error-related brain responses as a reward signal. <sup>49</sup>              | Implicit reactions are noisy and are not consent. See Part V.                     |
| <b>Co-adaptation</b>                          | No counterpart in language models: the person's brain adapts to the decoder while it adapts to them. <sup>38 39</sup> | Both sides are moving targets. Stability has to be designed.                      |
| <b>Distillation, on-device inference</b>      | A small personal model that runs locally.                                                                             | Wanted for privacy and delay; little published evidence yet for brain decoders.   |

Each cited result holds under its own study's conditions, named in the sources. The pairings and the catches are our synthesis.

One row deserves a closer look. An implanted handwriting decoder stayed accurate for 403 days without stopping its user to recalibrate. A language model corrected its output, and the corrections were fed back as training labels.<sup>35</sup> Others stay stable with no labels at all, by re-aligning each day's activity to the structure that persists beneath it.<sup>36 37</sup>

There is a trap in it. A decoder trained on its own corrected guesses learns from the language model as much as from the person. It can grow more fluent and less faithful at once. The cure is to anchor learning to what the person explicitly confirmed.

Deliberate confirmation is not only a safeguard. It is the cleanest training signal such a system will ever receive.

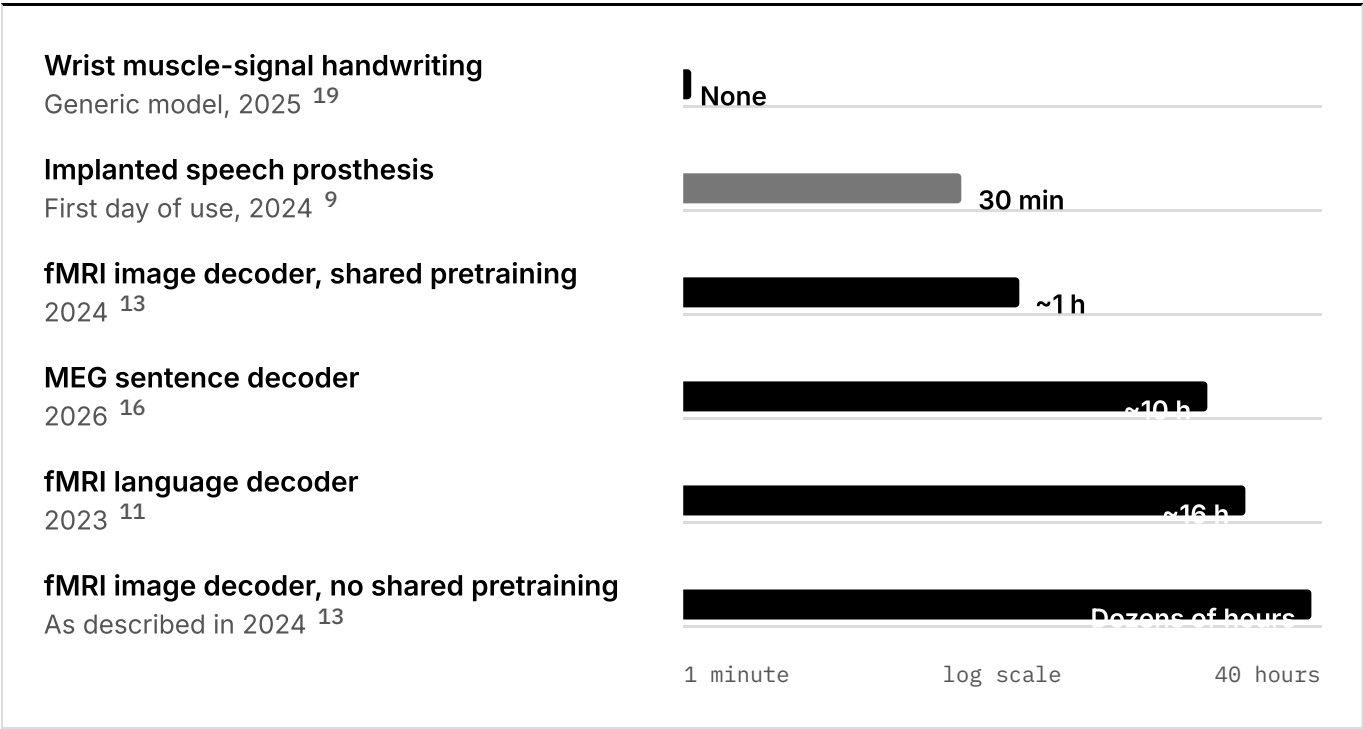
# How long before it knows you

One number says more about a system’s future than its headline accuracy: how much of a person’s time it needed before it worked.

FIGURE 12 | DATA

## How long before it knows you

The personal data each published system needed before it worked for a given person. The spread runs from nothing to dozens of hours.



**Personal recording time, logarithmic scale.** Figures are from each paper’s abstract except the fMRI language decoder, cited from the published record. Tasks, sensors and accuracy differ completely between rows, so this is not a ranking of systems; it shows only how widely the time demanded of a person varies. Outline = peripheral muscle signal, grey = implanted, black = brain recording without surgery.

This is why continuous adaptation is the way to build, and why training a model once and launching it is not enough. People differ. One person’s signals change from day to day. So the architecture has two parts: a large general model that knows what brain signals are like, and a small personal layer, fitted in minutes and updated continuously, that knows what yours are like.

If every user is fitted individually, nobody is the default user. Not one language, one head shape or one way of moving.

## The stranger problem

Train a decoder on one person on one day, then ask it to work on someone else, or on the same person next week. Choose a training cell and switch three remedies on and off.

- Works well
- Usable with corrections
- Unreliable
- ▨ Near chance

Trained on person C, day 3. It works there, fades across the same person's other days as sensors shift and the brain drifts, and fails on everyone else.

Anchors in the literature: a per-person fMRI language decoder did not transfer to other people;<sup>11</sup> shared pretraining cut new-person fMRI data to about an hour;<sup>13</sup> an implant recalibrated in minutes;<sup>9</sup> a wrist EMG model generalised with no calibration;<sup>19</sup> an implanted handwriting decoder recalibrated itself for 403 days.<sup>35</sup> Each is a different signal and task.

**Schematic.** Shading follows simple rules written by the author (same session > same person, later day > other person; pretraining and calibration each lift a cell by one level, never above "usable" for strangers; continual learning only helps the same person on later days). It is not fitted to any dataset. The cited studies show that each effect exists somewhere; none has yet shown them combined for EEG.

## Two learners

In one respect this is unlike the rest of machine learning. A language model's training data never learned anything back. Here the source of the data is a brain, and brains adapt. In animals, neural activity reorganised around a decoder as it was used.<sup>38</sup> In people, months of mutual training took two pilots with tetraplegia to an international competition for brain-controlled devices, with learning measured on both sides.<sup>39</sup>

The right picture is a musician and an instrument coming to fit each other. The system must be stable enough for the person to learn it and adaptive enough to follow the person as they change.

# Two loops

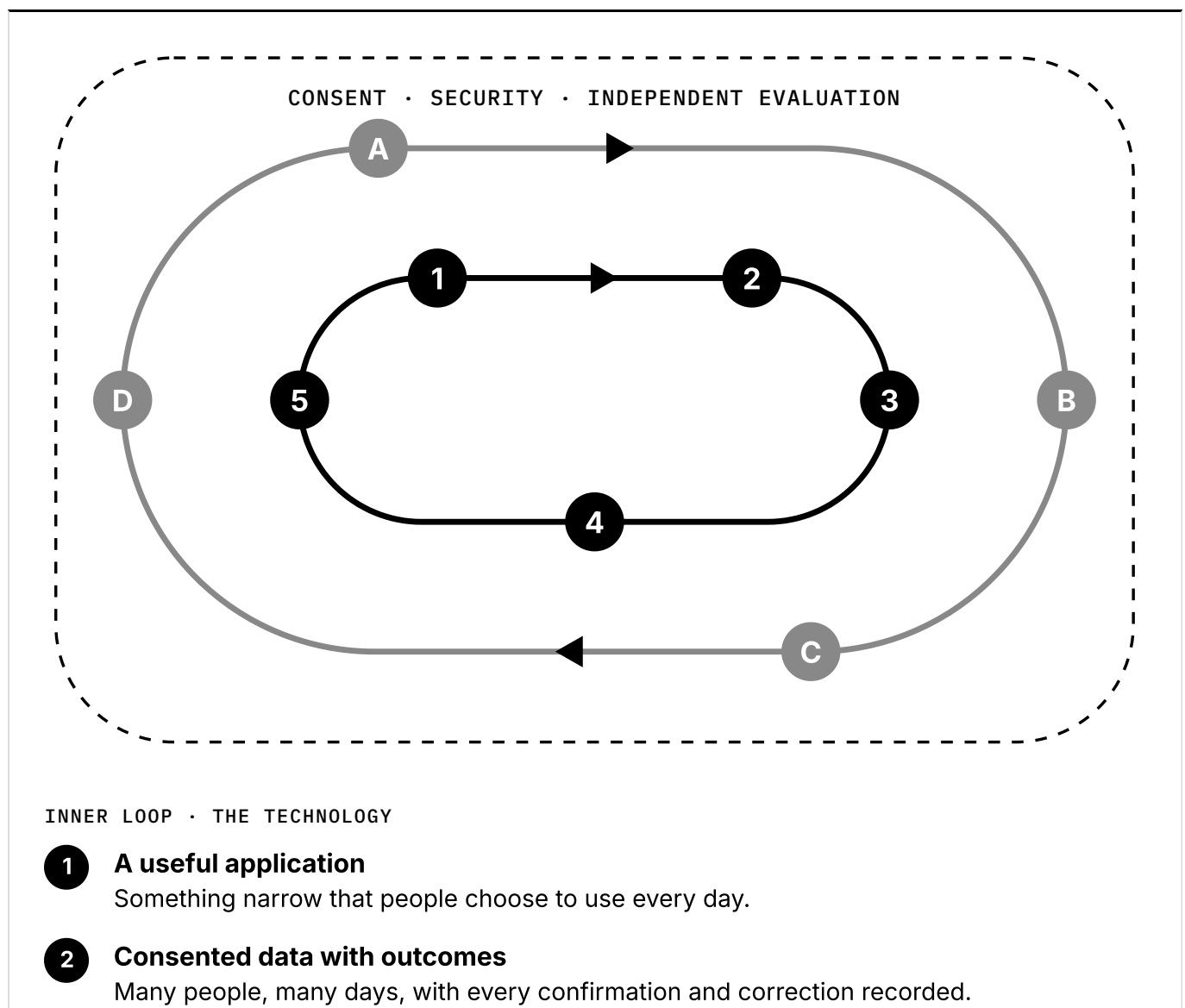
A device that is useful gets worn. Wearing produces data and, more valuable, outcomes: this decode was accepted, that one was corrected. Those improve the model, which shortens calibration, which makes the device useful to more people.

A second loop turns around the first. A product people keep wearing earns revenue. Revenue and results attract capital and talent. Capital and talent buy better sensors and more research. AI followed exactly this path. No field has scaled on grants alone.

FIGURE 14 | MODEL

## Two loops, one device

A device people choose to wear turns two wheels at once. The inner one makes the technology better. The outer one pays for it.



**3****Better models**

General representations that tolerate new heads and new days.

**4****Less calibration**

Minutes instead of hours to learn a new person.

**5****Wider use**

More people for whom it is worth wearing, and more applications.

**OUTER LOOP · THE MEANS****A****A product people keep wearing**

Retention is the first proof that the value is real.

**B****Revenue**

Sales that do not depend on the next grant or the next round.

**C****Capital and talent**

Results and revenue draw investors, engineers and scientists into the field.

**D****Better sensors and research**

Manufacturing at volume, lower cost per device, and funded science, feeding the next product.

**Conceptual model.** No quantities are implied. The dashed boundary is drawn around both loops on purpose: the fuel is recordings of people's nervous systems, and a loop that runs without consent runs on borrowed trust.

# The source

Text, code, images, audio and video are compressed products of human cognition. Neural data is the closest anyone can get to the engine that produced them. We think it becomes the next modality of AI.

Several fields are colliding at once. AI has become an engine of creation. Earlier technologies amplified human thinking and waited for it. This one increasingly does the iterating itself: it writes the code, designs the experiment and proposes what to try next.

So far this essay has treated BCI as a way of giving machines instructions. That is the smaller half. Neural data meets AI in three places: how models are trained, how people talk to them, and how they are kept aligned with what people want.

Two of those are needed today, by customers who already exist. Laboratories need new data, and they need a faster way to align their models. Neither has to wait for a consumer product. They are the immediate wedges.

## Five products and their source

Look at what AI has learned from. Text is compressed knowledge. Code is reasoning that can be run and checked. Images are perception. Audio is speech and timing. Video is motion. Every one is something a mind made, pressed into a format and stripped of nearly everything that produced it.

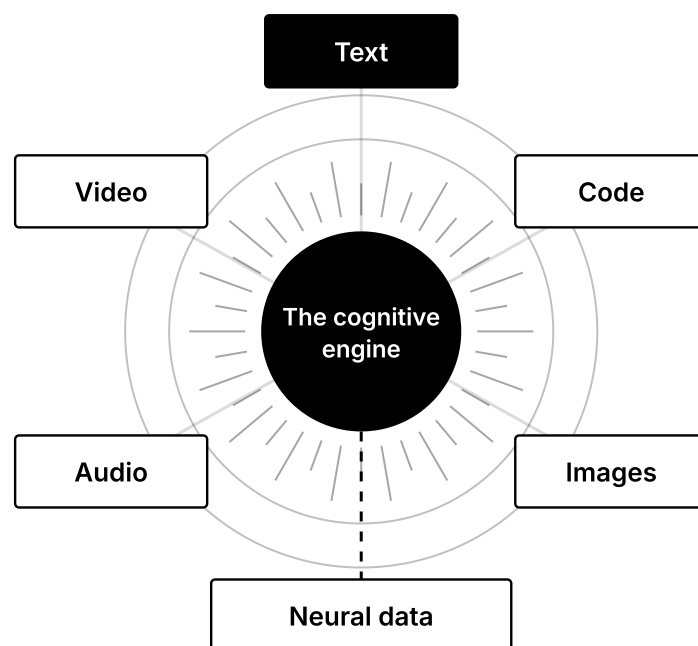
An essay keeps the conclusion and throws away the hesitations. A program keeps the logic and throws away the forty attempts before it worked. A photograph keeps the scene and loses what the photographer noticed in it.

Neural data is to those five what the Sun is to every other source of energy on Earth: the thing the rest are derived from.

FIGURE 15 | INTERACTIVE

## Five products, one source

Everything AI has learned from is a compressed product of human cognition. Neural data is the one recording taken at the engine itself. Select each to see what it keeps and what it loses.



### Text

|           |                                                                                       |
|-----------|---------------------------------------------------------------------------------------|
| PRESERVES | Conclusions, arguments and knowledge, in a deliberate, revised, error-corrected code. |
| LOSES     | The process: hesitations, discarded alternatives, what was noticed on the way.        |
| AVAILABLE | Effectively the public web.                                                           |

**The authors' thesis, drawn.** "Preserves" and "loses" are our characterisation. The dashed line marks the one connection that is still being built.

What lives at the source and never reaches the page is what makes experts expert: intuition, pattern recognition, the discomfort that says a result is wrong before you can say why, the hunch that comes before the proof. Philosophers call it tacit knowledge. We know more than we can tell. It cannot be scraped, because it was never written down.

The size of what is missing can be estimated. Behaviour leaves a person at roughly ten bits a second. The senses take in something like a billion.<sup>3</sup> Most of that gap is noise. But a person's state of mind lives somewhere inside it, and behaviour never carries it out.

Today's sensors capture a small fraction. We read that as the size of the opportunity. The source is there. The instrument is what has to improve.

## What the source contains

FIGURE 16 | TABLE

### What the source contains

Seven kinds of information in neural data that no finished text, image or recording of behaviour carries directly. Five are already measurable.

| SIGNAL              | WHAT HAS BEEN SHOWN                                                                                                                            | SENSOR                                 |
|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|
| Attention and focus | Which of two simultaneous speakers a listener is following, from about a minute of unaveraged signal. <sup>50</sup>                            | Scalp EEG                              |
| Error detection     | A response within a fraction of a second of seeing a mistake, detectable on single trials and already used to teach a device. <sup>48 49</sup> | Scalp EEG                              |
| Intent              | Intended movement decoded well enough to drive cursors, robotic arms, handwriting and speech. <sup>6 20</sup>                                  | Implants; scalp EEG at lower dimension |
| Semantics           | The gist of heard language and the look of viewed images; <sup>11 12</sup> semantic maps consistent across people. <sup>56</sup>               | fMRI, MEG                              |
| Language and syntax | Typed and attempted sentences decoded word by word. <sup>16 8</sup>                                                                            | MEG; implants                          |

| SIGNAL                         | WHAT HAS BEEN SHOWN                                                                                                                                                         | SENSOR               |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| <b>Preference and judgment</b> | Value and affect signals reported in controlled studies. Making them robust for one person across days is the engineering frontier.                                         | Scalp EEG and others |
| <b>Reasoning</b>               | The path to a conclusion: intuition, pattern recognition, discomfort, the hunch. Not yet recorded in recoverable form. We expect synchronised data at scale to change that. | To be built          |

Cited rows summarise those studies under their own conditions. The last two rows state where we expect the field to go.

Notice which sensor does the reading. Attention and error detection are read with EEG today, and so are the rhythms of sleep.<sup>50 48 51</sup> Meaning is laid out across the cortex in patterns consistent enough between people to draw an atlas.<sup>56</sup>

Three properties make this data unusually valuable. It is raw: nobody has decided in advance what mattered, so the model can. It is synchronised: a recording time-locked to what the person saw, did and corrected labels every moment of a task for free. And it is personal: each user gets a small model that maps their brain into the same shared space that language and vision models already use.

## Brains have already taught models

Artificial neural networks began as abstractions of biological ones. AI then scaled far past its source. The flow is now running back.

Language models trained only to predict the next word turn out to predict human brain responses to sentences, and the better the model, the closer the match.<sup>44 45</sup> Two different systems, set the same problem, converged. And when brain data is put into training, models improve. Image classifiers nudged toward the organisation of a mouse’s visual cortex became more robust.<sup>46</sup> Speech models tuned on brain recordings of people listening to stories became better at meaning.<sup>47</sup>

Those gains came from a few animals and a few listeners. Nobody has run the experiment at scale. That is the experiment worth running.

## The first wedge: data for AI labs

The old fuel is running out. On current trends, models will have consumed essentially all public human text by around 2028.<sup>43</sup> The frontier laboratories know it, which is why they now pay for data they once scraped for nothing.

Every new kind of data they added bought a new capability. Code brought reasoning that can be checked. Images brought perception. Audio and video brought speech and motion. The next kind has to come from somewhere, and neural data is the one modality nobody has. It cannot be scraped and it cannot be synthesised. Someone has to build the sensors and gather it.

Our working estimate: at about \$300 per recorded hour, ten million hours of neural recording costs roughly three billion dollars. That is less than one large GPU cluster, in an industry planning to spend hundreds of billions on compute. And the data does not need to be perfect. At the valuations frontier models now command, a corpus that improved reasoning or alignment by one percent would repay itself many times over.

That is why this wedge is immediate. A hardware company normally has to wait for millions of consumers before its data is worth anything. Here the buyer exists today, and has deeper pockets than any consumer market. A laboratory can fund the corpus before the consumer product is finished. Better interfaces will produce better models, and better models will fund better interfaces.

That corpus cannot be recorded in hospitals. It can only come from an inexpensive device worn by very large numbers of consenting people. In practice that means EEG.

We expect a frontier laboratory to name neural data in the training of a flagship model within a few years.

# The lossy postcard

A prompt is a postcard from a rich internal state. It leaves out the tone you wanted, the constraint you forgot to mention, the thing the last draft got wrong. Models fail on that compression at least as often as they fail on intelligence.

A BCI widens the channel twice. It lets intent out with less effort. And it lets context in without being asked for: what you are attending to, whether you are lost, when you have noticed a mistake. In time, prompting becomes something that happens continuously in the background.

The same channel speeds the making of AI. Labelling data is a person looking at examples and clicking. A reader whose recognition and doubt can be sensed directly could judge a stream of examples many times faster.

## The second wedge: alignment

The biggest point may be this one.

AI systems are shaped after training by human judgment: raters comparing answers, users pressing approve or try again. This is how models are aligned with what people want, and a human can judge only so many answers in an hour. Feedback is the scarcest ingredient in making models behave well, and every laboratory needs more of it, for every model, now.

Consider the error signal: an EEG response that says *that was wrong* within a few hundred milliseconds, without the person lifting a finger. It has already been used as a teacher. A device tries actions, the person only watches, and the error responses steer the device toward the goal.<sup>49</sup>

Extend that from a robot to a model and neural signals become an alignment technique: a literal, better RLHF. A click carries one bit, yes or no, after the fact. A reaction carries when the doubt began, which sentence caused it and how strong it was, as it happens. The human stops being the bottleneck in the loop and becomes its richest signal.

This needs no consumer product either. A laboratory can put the sensor on its own raters tomorrow. The first users of an alignment interface are the people already paid to judge

models, which is what makes this the second immediate wedge.

It has one requirement. A reaction is not a decision. Systems trained on clicks learned to capture attention, and nobody should repeat that with a deeper probe. Reactions may inform a system. Only decisions should authorise it.

## From watching the machine to joining it

There are two ways to keep advanced AI in line with human interests. The first works from outside: rules, evaluation, governance. It is indispensable now, and it weakens as machines outpace the people watching them. The second works from inside, by coupling human intent to the system closely enough that steering does not have to be imposed afterward. BCI belongs to the second, and almost nothing else does.

It contributes in three ways. A field that decodes the brain at scale produces the best available map of how values are held in a mind, which is guidance for building systems meant to share them. A richer channel shows what a click cannot: whether the person understood the plan, or was merely soothed. And over a longer horizon, AI becomes less a separate agent to be watched and more an extension of a person's own thinking.

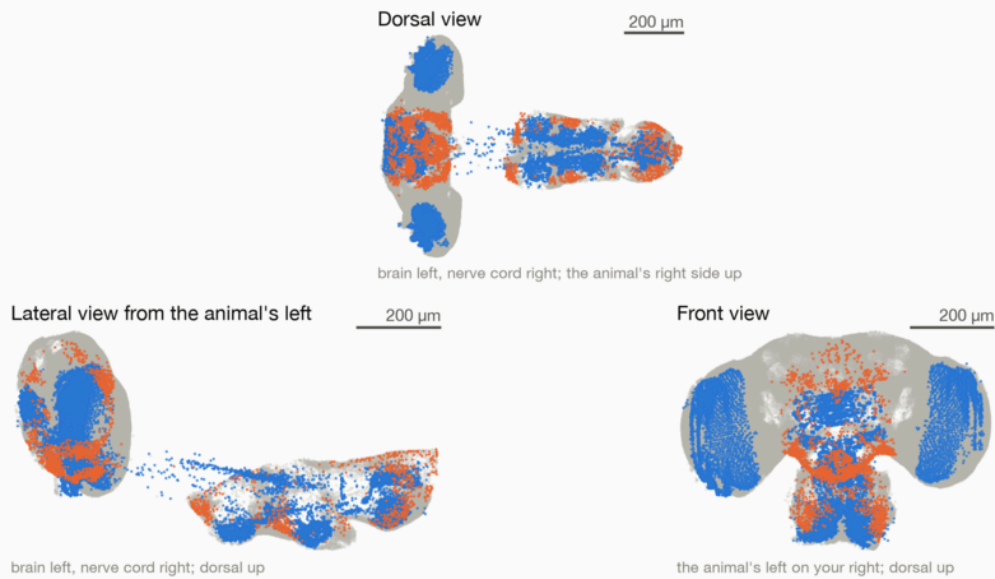
One question hangs over all of it: aligned to whom? A system tuned to the minds of one company or one state would be a triumph of engineering and a failure of everything else. The signal has to come from many people, over long periods, and keep changing as people do.

## A graph, and how graphs get solved

The brain is a graph, and the task is to make meaning out of it. The complete brain of a fruit fly has been mapped: 139,255 neurons and fifty million synapses.<sup>55</sup> A human brain has about 86 billion neurons.<sup>60</sup> Nobody will trace that by hand.

**Where the 17,937 sensory input neurons sit against the 2,131 descending and motor readout neurons; the other neurons are the fixed wiring between them**

● Sensory input 17,937    ● Descending 1,316 and motor 815 readout    ● All other neurons 146,632



Positions: soma location 139,662; tracing-to-soma location 976; centroid of the neuron's own synapses 25,941 (mostly sensory neurons, cell bodies outside the volume); not drawn 121. 8 nm voxels converted to micrometres. Positions from body-annotations-male-cns-v1.0-minconf-0.5.feather (soma) and syn-points-male-cns-v1.0-minconf-0.5.feather (synapses). Every neuron is drawn at the same marker size.  
Source: MaleCNS v1.0 (HHMI Janelia, Cambridge, MRC LMB, Google Research), CC BY 4.0, released 8 Jun 2026; edges  $\geq 5$  synapses; built 20 Sep 2026.

Every neuron in a fruit fly's central nervous system, positioned from an electron-microscope reconstruction. A human brain has hundreds of thousands of times as many. Image: Jason Alan Snyder; data: the MaleCNS team, CC BY 4.0.

Nobody traced the internet by hand either. The web is the largest graph humans ever built, and its meaning was extracted by statistics at scale, by models fed enough of it. That is how a graph too large to read gets solved, and it is the template for the brain.

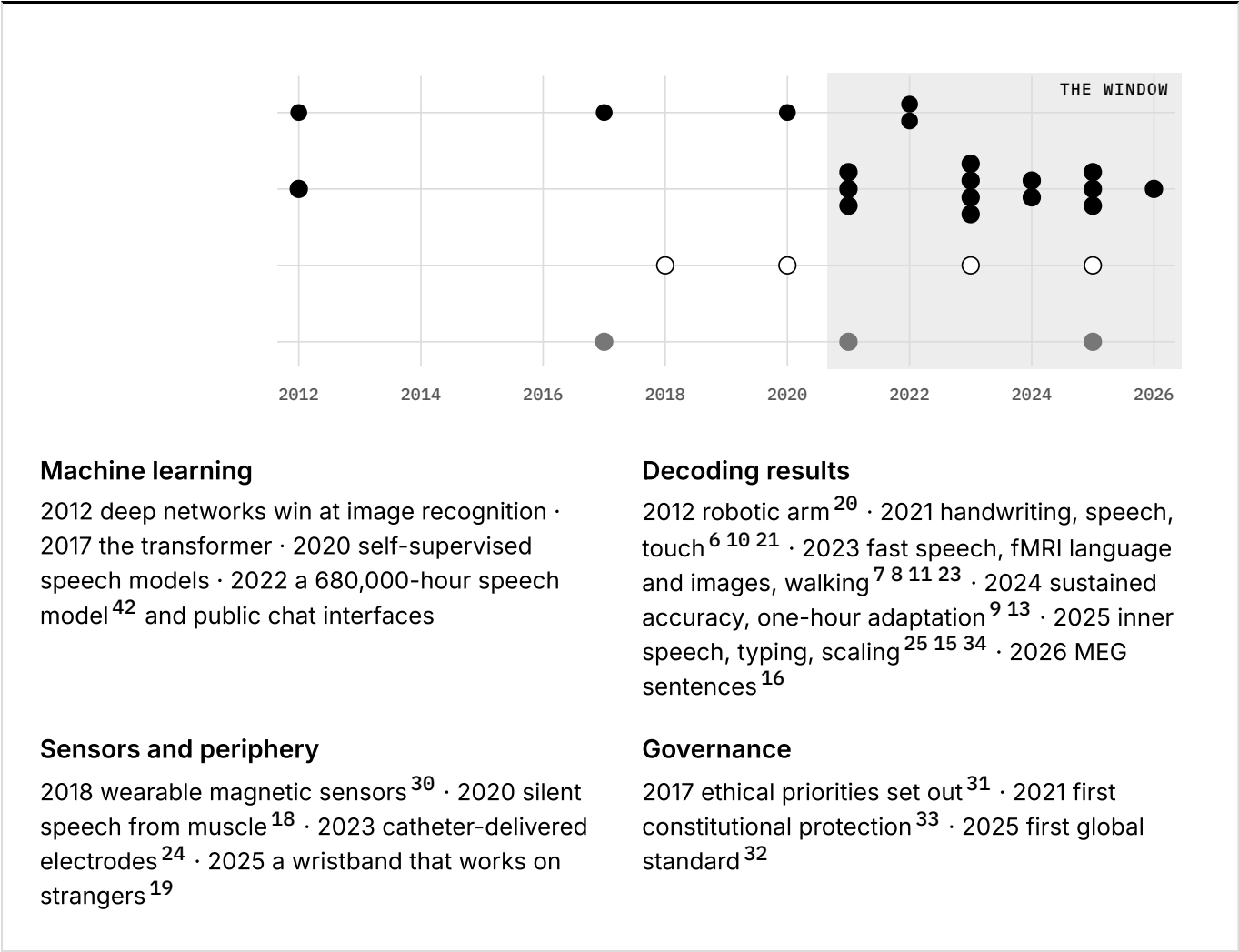
## Why now

Because everything that was missing has arrived inside the same few years.

FIGURE 17 | TIMELINE

## Four clocks striking together

The milestones this essay relies on, by year. Models, decoding results, wearable sensing and governance each matured separately. Most of the dots fall after 2020.



**A selected record, not a census.** Each dot is a milestone cited in this essay, placed at its publication year; the three earliest machine-learning milestones are common knowledge and uncited. Because the essay concentrates on recent work, some of the clustering reflects that choice. The figure shows coincidence in time and makes no claim about what comes next.

The models exist. The proofs exist. EEG hardware has become cheap, and a decoder fits in a pocket. The demand exists, because capable agents have raised the value of every instruction a person gives. And the old data is nearly spent.

That is a window, and it has two edges. The groups that build the sensors and gather the first large, consented datasets will set the direction of the field. And the terms on which people and machines relate are being set now, on a schedule of months. An interface that

arrives after those terms harden arrives too late to shape them. Surgery cannot scale inside that interval. EEG can.

# The road from here

BCI reaches the world through the first product millions of people choose to wear. This is the order we expect, and what Setu is building.

## Where it shows up first

Applications will arrive in order of how much signal each one needs. Each rung is a step up in signal quality, and each unlocks a product that was impossible on the rung below.

FIGURE 18 | SEQUENCE

## Seven steps up, and two beyond

The order in which we expect non-invasive interfaces to become useful, from the least signal required to the most, with what already exists at each step.

**1 Emotional and psychological biomarkers**  
Coarse state is what today's wearables already measure. Sleep and stress come first; controlled evidence for therapy is still maturing.<sup>54</sup>

**3 Subvocal text**  
Muscle signals decoded into text across people with no calibration.<sup>19</sup>

**5 Visual reconstruction of images**  
Shown in scanners,<sup>12 14</sup> with scaling measured down to EEG.<sup>34</sup>

**2 Attention and focus tracking**  
The attended speaker can already be read from a minute of scalp EEG.<sup>50</sup>

**4 Silent speech**  
Mouthed words turned into audible speech from facial muscle signals.<sup>18</sup>

**6 Multimodal decoding**  
Several signals and context fused into one estimate of intent.

7

**Continuous transcription**

Sentences from MEG during typing;<sup>16</sup> inner speech from implants.<sup>25</sup> The summit of the climb.

8

**Replay of remembered scenes**

Beyond the seven: an imagined or remembered scene with its sound and motion.

9

**A two-way channel**

Reading and writing together, where “what I thought” and “what we worked out” become hard to separate.

**A proposed sequence.** The ordering is the authors’ expectation. The note under each step cites published work under its own conditions, which in several cases means a laboratory, a scanner or an implant and not yet a wearable.

A second ladder runs beside the first. It measures how much a new person must give before the device works for them. Today that is hours in a laboratory. Below it: a decoder that survives from one day to the next, a model pretrained across many people, adaptation from a few seconds of use, and finally nothing at all. A consumer will not train a device for an afternoon. They will put it in a drawer.

## The wedge

Part V named the customers who need this now: laboratories buying data and laboratories aligning models. They fund the beginning. Reaching everyone takes a product people choose to wear.

People do not buy scaling laws. They buy something that solves a problem or feels like magic. We think two products get the first ten million devices onto ten million heads.

**Communication.** Silent, private, hands-free exchange with machines and, through them, with each other. It is the nearest thing to telepathy that engineering can honestly offer. It does not need to beat a keyboard. It needs to work where a keyboard and a voice cannot: on a train, in a meeting, with hands full, wearing glasses that have no keyboard at all. An AI assistant that acts for you all day is of little use if you can only address it aloud in an empty room.

**Brain wellness and recovery.** People already wear rings and bands that score how the body recovered overnight. The brain is the missing reading, and it is an EEG reading. EEG

has its strongest clinical evidence here. When a stroke survivor's attempt to move a paralysed arm is detected in their EEG and answered at once with stimulation of the muscle, the brain rewires: patients recovered more function than with sham treatment and kept it six to twelve months later.<sup>52 53</sup> People will wear something that helps them sleep, focus and recover long before they wear something that types.

What makes someone wear either one every day is everything around the decoder: a form that disappears into glasses, earbuds or a band, no calibration ritual, a battery that lasts the day, a price nearer to headphones than to a medical device, and a benefit felt in the first week.

## What Setu is building

Setu exists to redefine how people and machines communicate, without surgery. We build the full stack, from sensor to model to product.

Our destination is EEG: a brain interface cheap and comfortable enough for anyone to wear. We start one rung below it on purpose. Silent speech, read from the muscles and fine movements of the face and throat, is where the signal is already strong enough to make something people can use today, and assistive communication is the first place it changes a life.

We are plain about the difference. A muscle sensor is not a brain sensor, and success at one does not prove the other. What carries across is everything else: wearable engineering, interaction design, models that keep learning each user, and people who trust the maker. The first product pays for the climb and teaches us how to build something a person will keep on.

## Beyond reading

This essay has stayed in one corner of a large map: interfaces, without surgery, that read. That corner is where scale begins. The capital, talent and methods it produces flow into the rest of neurotechnology.

**Writing.** Reading exploits a leak: brain activity casts faint shadows through the skull. There is no such leak in reverse. Placing a precise pattern into chosen neurons from outside the head is far harder. Coarse forms are already in clinical use, and we expect the rest to accelerate as reading draws people into the field.

**Implants.** The highest fidelity will always be inside the skull. It will also arrive at scale last. What EEG builds, the implant makers inherit.

**Cognition.** A perfect interface still leaves a biological brain thinking at biological speed. Augmenting the brain itself, or emulating one in a computer, is the far horizon that addresses that.

## Three things worth building

---

### **Brain-aware environments**

Software that is no longer blind to the person using it: an assistant that answers briefly when you are tired and at length when you are confused, a lesson that changes course the moment attention drifts.

---

### **The inner space stack**

Tools for a person's own mind: meditation that measures depth instead of minutes, sleep that is shaped as well as tracked, mental-health support that responds to what is happening now.

---

### **The creative engine**

Holding an image, a melody or a structure in mind and watching a first draft appear. The machine supplies the range. The person supplies the taste.

And, eventually, communication between minds. A link from one brain to another is two interfaces back to back: one that reads and one that writes. It will arrive through machines first, with each person's interface translating into and out of language.

# The decade ahead

Deep changes arrive the way water freezes: a long, invisible fall in temperature, then slush, then ground you can stand on.

**Quiet.** In the next few years, the first general models for brain data become usable outside the laboratories that made them. The first devices a healthy adult would choose to keep on begin to ship. Engineers who spent a decade on language models start moving to the interface.

**Loud.** Around the turn of the decade, BCI has its ChatGPT moment. One device passes ten million regular users. Silent input starts to displace voice and thumbs. The hardware disappears into glasses, earbuds and headbands. Mental privacy becomes an ordinary political argument.

**Ordinary.** In the 2030s the technology stops needing explanation. Robots take direction through the same channel as software. Writing to the brain becomes practical for the first time, and with it come the hardest questions in this essay.

One morning from that third period. Ines coordinates a hospital ward. On her way in she reviews the night's handover on her glasses, and each time something looks wrong to her, the item is flagged before she has decided to flag it. She composes three messages without speaking, reads each one back, and sends two. At the ward she asks the logistics agent, silently, to move a delivery. It tells her what that will delay, and waits. She thinks: no, leave it. It leaves it. None of this feels remarkable to her. It feels like having her hands back.

# A layer you can switch off

The measure of this technology is not how much of a person becomes legible. It is how much a person can do, and decline to do.

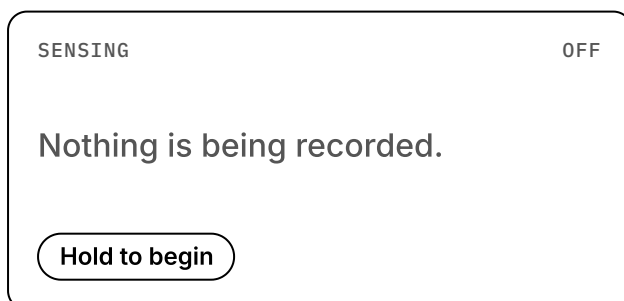
For as long as there have been people, the inside of the skull has been the one room nobody else could enter. A technology that reads from it has to earn the right to exist.

The near versions do not require reading a mind. They require a system that can tell a deliberate act from everything else a nervous system does, and that treats the difference as sacred. That system is the intent layer: sensing, interpretation and learning, behind a gate only the user can open.

FIGURE 19 | STORYBOARD

## An interface that asks first

Six frames of a concept interaction. Nothing is sensed until the user opens the gate, nothing is sent until they have seen it, and everything can be taken back.



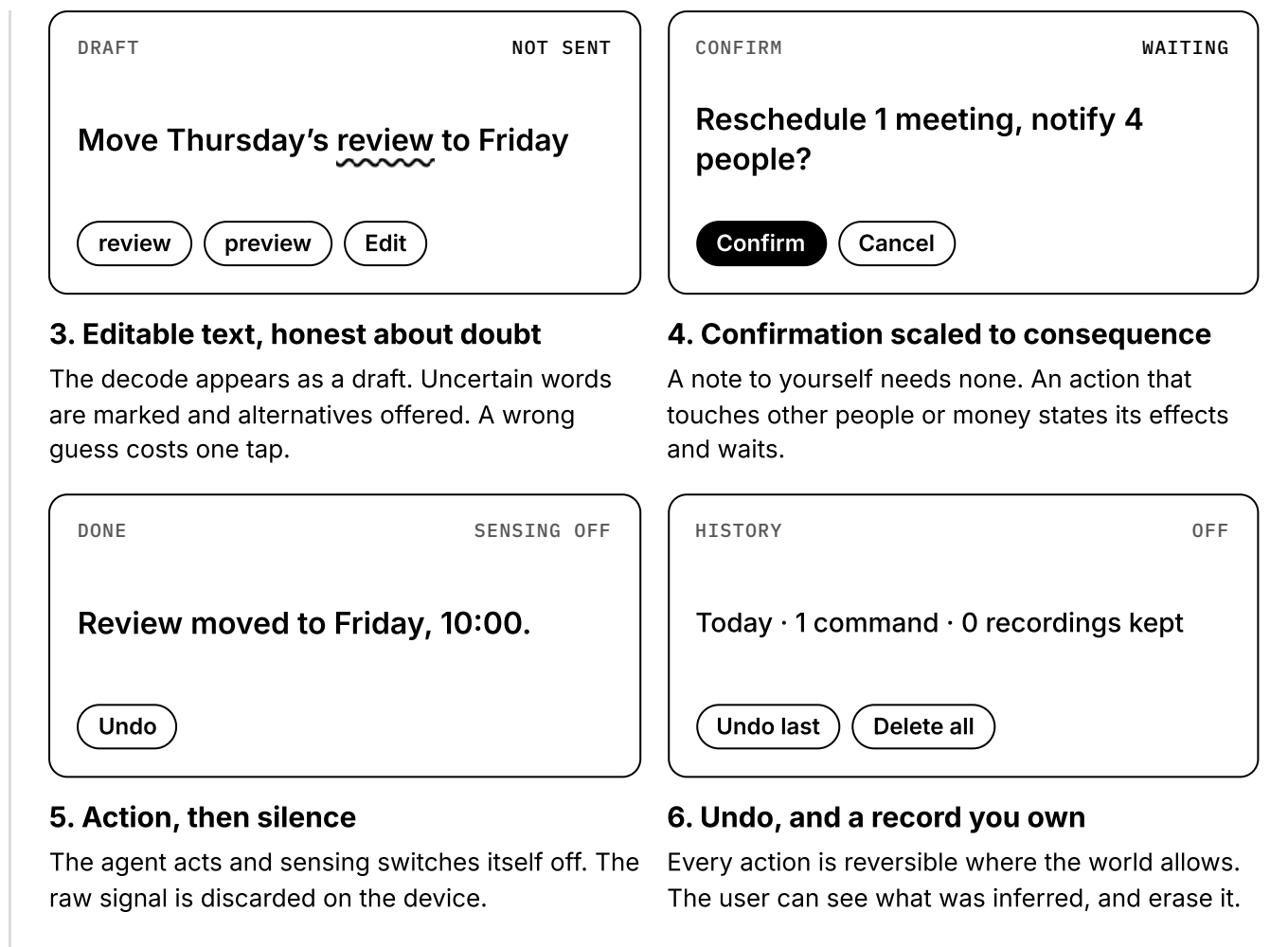
### 1. Off by default

The sensor is idle until a deliberate, physical option. Its state is always visible to the wearer and to people nearby.



### 2. A deliberate signal

The user makes a specific, intended act: mouthing words, or attempting to speak. Idle activity is ignored and never stored.



**Concept storyboard.** This is a design principle drawn as an interface, not a depiction of a working product or of any measured performance. The gate in frame 1 has laboratory precedent: implanted speech decoders have been shown to withhold inner speech unless the user deliberately unlocks decoding.<sup>25</sup>

## Who holds the interface

BCI began as a clinical technology. It is becoming a productivity technology. It will become a data technology once neural recordings count as training material, and then the ordinary channel between people and machine intelligence. Its strategic weight rises at each step.

States have noticed. One major power names BCI among its strategic future industries. Another leads in startups and frontier AI. A third is writing the rules for mental privacy that others will have to meet. Defence agencies have funded the field for decades. The AI decade was a contest over chips. The next will include the sensors that record the nervous system, and the question of who may collect, model and act on what they record.

We do not think this should be a technology a few countries own. The people with most to gain from a cheap interface live everywhere and speak every language. Open standards and devices priced for ordinary incomes are how it reaches them.

## What can go wrong

- **A recording says more later than it does now.** A trace that shows only drowsiness to this year's models may show mood or health to the models of five years on. Consent given today covers inferences nobody can yet list.
- **Manipulation.** A system that can sense a reaction can learn to produce one.
- **Drift of the self.** Did I choose that, or was I completed?
- **Intrusion.** Everything connected can be attacked, and this is the most intimate attack surface ever shipped.
- **Surveillance.** Monitoring attention at work or school, profiling dissent, interrogation. Someone will propose each.
- **Inequality.** An interface to machine intelligence owned by a few and priced for a few.

The state of practice is not reassuring. A review of thirty consumer neurotechnology companies found that nearly all reserved broad rights over the neural data they collect.<sup>62</sup>

## How this goes right

Good outcomes have to be designed, and the strongest safeguards are technical before they are legal. A law can be repealed. An architecture has to be rebuilt.

- **Process on the device.** Send intent, not waveforms.
- **Keep neural data away from advertising.** Always.
- **Make consent expire.** Next year's models will read what this year's cannot.
- **Give neural data its own legal category.** No resale, no hidden inference, no access for employers or insurers.
- **Hold writing to a higher bar than reading.**

- **Agree the red lines early.** No neural interrogation, no compulsory monitoring, no covert stimulation.

The work has started. Neuroscientists set out the case for mental privacy in 2017.<sup>31</sup> One country has written protection of brain activity into its constitution.<sup>33</sup> In November 2025 the member states of UNESCO adopted the first global standard.<sup>32</sup> For once, rules may arrive before the scandal.

There is a failure on the other side too. Regulate the field to a standstill and the result is not safety. It is a frozen technology in the years it matters most.

**What Setu holds itself to.** Sensing is off unless the user turns it on. Raw signal stays on the device wherever the engineering allows. We decode what was deliberately offered and nothing else. Every output can be seen, corrected and undone before it acts. Neural data will not be used for advertising. Nobody should have to wear our product as a condition of a job or a classroom.

## The case against us

---

### OBJECTION

**The keyboard has survived fifty years of predictions. Hands and voices are superb.**

They are, and they did not evolve to direct machines a thousand times faster than they can type. BCI starts where hands and voice fail and grows from there.

---

### OBJECTION

**These results come from single patients in heroic laboratory conditions.**

Correct. One participant is an existence proof. It shows the information is there. Turning it into something a stranger can wear is the rest of this essay.

---

### OBJECTION

**The skull is a filter. No dataset repairs physics.**

The skull sets a ceiling somewhere. Nobody has found it. Every EEG curve measured so far is still rising. Speech recognition was not forbidden by physics either. It needed a hundred thousand hours.

---

OBJECTION

**Effort spent outside the skull is taken from implants, which have the higher ceiling.**

For many years only a device without surgery can reach millions. The methods, the manufacturing skill and the people it produces carry across.

---

OBJECTION

**The brain is not the internet. There is no objective as clean as predicting the next word.**

The objective is prediction across signal and context: given a recording and what the person saw, heard and did, predict the missing part. AI has scaled on messier objectives.

---

OBJECTION

**Consumer EEG headsets have been tried. The category is a graveyard.**

The earlier wave built products before there was a stack beneath them: no general models, no buyer for the data, no agents that made a new input valuable. Most technologies fail until the missing pieces exist.

---

OBJECTION

**It is too private ever to be collected at scale.**

That is why the safeguards have to be in the first device. A version that treats neural data the way advertising treats browsing history deserves to fail.

---

OBJECTION

**The channel could strengthen the machine more than the person.**

It could. Nothing about the technology favours the human automatically. That has to be designed.

---

## What would prove us wrong

Our case rests on three assumptions. EEG carries enough information. Decoding keeps improving as data grows to a hundred thousand hours and to a million. And models trained across enough people cut each new person's calibration from hours to minutes.

Anyone is free to disagree, and the disagreement is testable. If any of the three fails, the strong version of this thesis is wrong. We do not expect that.

## The condition

The cheaper and more widely worn a sensor becomes, the more people it helps and the more intimate data exists in the world. Wider access and deeper exposure arrive together. Good design decides that tension in the user's favour, every time, by default.

Beneath the policy sits a choice no regulator can make for us: what we want to keep about being human. Agency. Authorship. The privacy of an inner life.

This future will not be good by default. It has to be built on purpose, and not by some later generation. If you build, fund, regulate or simply care how people will live beside intelligent machines, find where you can push. We would like to hear from you.

If it is built, it will be judged by two silences. The first is the one it ends: the silence of a person who has something to say and no way to say it. The second is the one it must keep: the silence of everything a person chose not to say.

An interface that ends the first and guards the second would be a new layer of computing. One that ended both would be something else, and we will not build it.

---

# Sources and further reading

Quantitative claims were checked against the abstracts of the primary publications in October 2026. Preprints are marked and have not completed peer review. Entries 1, 17, 30, 59, 60, 61 and 62 are cited from the published record without being re-fetched during that check.

1. Dhakal V, Feit AM, Kristensson PO, Oulasvirta A. Observations on typing from 136 million keystrokes. *Proc. CHI* (2018). doi:10.1145/3173574.3174220
2. Coupé C, Oh YM, Dediu D, Pellegrino F. Different languages, similar encoding efficiency. *Science Advances* 5, eaaw2594 (2019). doi:10.1126/sciadv.aaw2594
3. Zheng J, Meister M. The unbearable slowness of being: why do we live at 10 bits/s? *Neuron* 113 (2025; online 2024). doi:10.1016/j.neuron.2024.11.008
4. Wolpaw JR et al. Brain–computer interfaces for communication and control. *Clinical Neurophysiology* 113, 767–791 (2002). doi:10.1016/S1388-2457(02)00057-3
5. Vidal JJ. Toward direct brain–computer communication. *Annual Review of Biophysics and Bioengineering* 2, 157–180 (1973). doi:10.1146/annurev.bb.02.060173.001105
6. Willett FR et al. High-performance brain-to-text communication via handwriting. *Nature* 593, 249–254 (2021). nature.com/articles/s41586-021-03506-2
7. Metzger SL et al. A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* 620, 1037–1046 (2023). nature.com/articles/s41586-023-06443-4
8. Willett FR et al. A high-performance speech neuroprosthesis. *Nature* 620, 1031–1036 (2023). nature.com/articles/s41586-023-06377-x
9. Card NS et al. An accurate and rapidly calibrating speech neuroprosthesis. *New England Journal of Medicine* 391, 609–618 (2024). doi:10.1056/NEJMoa2314132
10. Moses DA et al. Neuroprosthesis for decoding speech in a paralyzed person with anarthria. *New England Journal of Medicine* 385, 217–227 (2021). doi:10.1056/NEJMoa2027540
11. Tang J, LeBel A, Jain S, Huth AG. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience* 26, 858–866 (2023). nature.com/articles/s41593-023-01304-9
12. Scotti PS et al. Reconstructing the mind’s eye: fMRI-to-image with contrastive learning and diffusion priors. *NeurIPS* (2023). arXiv:2305.18274
13. Scotti PS et al. MindEye2: shared-subject models enable fMRI-to-image with 1 hour of data. *ICML, PMLR* 235 (2024). proceedings.mlr.press/v235/scotti24a

14. Gong Z et al. NeuroClips: towards high-fidelity and smooth fMRI-to-video reconstruction. *NeurIPS* (2024, oral). arXiv:2410.19452
15. Lévy J et al. Brain-to-text decoding: a non-invasive approach via typing. Preprint (2025). arXiv:2502.17480
16. Zhang M et al. Accurate decoding of natural sentences from non-invasive brain recordings. Preprint (June 2026). ai.meta.com/research/publications
17. Défossez A et al. Decoding speech perception from non-invasive brain recordings. *Nature Machine Intelligence* 5, 1097–1107 (2023). doi:10.1038/s42256-023-00714-5
18. Gaddy D, Klein D. Digital voicing of silent speech. *Proc. EMNLP* (2020). aclanthology.org/2020.emnlp-main.445
19. Kaifosh P, Reardon TR, CTRL-labs at Reality Labs. A generic non-invasive neuromotor interface for human–computer interaction. *Nature* 645 (2025). doi:10.1038/s41586-025-09255-w
20. Hochberg LR et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485, 372–375 (2012). doi:10.1038/nature11076
21. Flesher SN et al. A brain–computer interface that evokes tactile sensations improves robotic arm control. *Science* 372, 831–836 (2021). doi:10.1126/science.abd0380
22. Edelman BJ et al. Noninvasive neuroimaging enhances continuous neural tracking for robotic device control. *Science Robotics* 4, eaaw6844 (2019). doi:10.1126/scirobotics.aaw6844
23. Lorach H et al. Walking naturally after spinal cord injury using a brain–spine interface. *Nature* 618, 126–133 (2023). doi:10.1038/s41586-023-06094-5
24. Mitchell P et al. Assessment of safety of a fully implanted endovascular brain–computer interface for severe paralysis in 4 patients. *JAMA Neurology* 80, 270–278 (2023). doi:10.1001/jamaneurol.2022.4847
25. Kunz EM et al. Inner speech in motor cortex and implications for speech neuroprostheses. *Cell* 188 (2025). doi:10.1016/j.cell.2025.06.015
26. Kuruppu G et al. EEG foundation models: a critical review of current progress and future directions. Preprint (2025). arXiv:2507.11783
27. Liu D et al. EEG-FM-Compass: progress, benchmarking, and future directions for EEG foundation models. Preprint (2026); journal reference *National Science Review*. arXiv:2601.17883
28. Nunes de Oliveira JM et al. Scaling laws for EEG decoding: how much data is enough? Preprint (September 2026). arXiv:2609.35056
29. Allen EJ et al. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience* 25, 116–126 (2022). doi:10.1038/s41593-021-00962-x
30. Boto E et al. Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* 555, 657–661 (2018). doi:10.1038/nature26147
31. Yuste R et al. Four ethical priorities for neurotechnologies and AI. *Nature* 551, 159–163 (2017). doi:10.1038/551159a

32. UNESCO. Recommendation on the Ethics of Neurotechnology. Adopted by the General Conference, November 2025. [unesco.org/en/ethics-neurotech](https://unesco.org/en/ethics-neurotech)
33. Republic of Chile. Law 21.383, amending Article 19 of the Constitution to protect mental integrity and brain activity in the face of neurotechnology (2021). [bcn.cl/leychile](https://bcn.cl/leychile)
34. Banville H, Benchetrit Y, d'Ascoli S, Rapin J, King J-R. Scaling laws for decoding images from brain activity. Preprint (2025). [arXiv:2501.15322](https://arxiv.org/abs/2501.15322)
35. Fan C et al. Plug-and-play stability for intracortical brain–computer interfaces: a one-year demonstration of seamless brain-to-text communication. *NeurIPS* 36, 42258–42270 (2023). [arXiv:2311.03611](https://arxiv.org/abs/2311.03611)
36. Karpowicz BM et al. Stabilizing brain–computer interfaces through alignment of latent dynamics. *Nature Communications* 16, 4662 (2025). doi:10.1038/s41467-025-59652-y
37. Degenhart AD et al. Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. *Nature Biomedical Engineering* 4, 672–685 (2020). doi:10.1038/s41551-020-0542-9
38. Orsborn AL et al. Closed-loop decoder adaptation shapes neural plasticity for skillful neuroprosthetic control. *Neuron* 82, 1380–1393 (2014). doi:10.1016/j.neuron.2014.04.048
39. Perdakis S, Tonin L, Saeedi S, Schneider C, Millán JdR. The Cybathlon BCI race: successful longitudinal mutual learning with two tetraplegic users. *PLoS Biology* 16, e2003787 (2018). doi:10.1371/journal.pbio.2003787
40. Azabou M et al. A unified, scalable framework for neural population decoding. *NeurIPS* (2023). [arXiv:2310.16046](https://arxiv.org/abs/2310.16046)
41. Jiang W-B, Zhao L-M, Lu B-L. Large brain model for learning generic representations with tremendous EEG data in BCI. *ICLR* (2024). [arXiv:2405.18765](https://arxiv.org/abs/2405.18765)
42. Radford A et al. Robust speech recognition via large-scale weak supervision. Preprint (2022). [arXiv:2212.04356](https://arxiv.org/abs/2212.04356)
43. Villalobos P et al. Will we run out of data? Limits of LLM scaling based on human-generated data. Preprint (2022, revised 2024). [arXiv:2211.04325](https://arxiv.org/abs/2211.04325)
44. Schrimpf M et al. The neural architecture of language: integrative modeling converges on predictive processing. *PNAS* 118, e2105646118 (2021). doi:10.1073/pnas.2105646118
45. Goldstein A et al. Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience* 25, 369–380 (2022). doi:10.1038/s41593-022-01026-4
46. Li Z et al. Learning from brains how to regularize machines. *NeurIPS* (2019). [arXiv:1911.05072](https://arxiv.org/abs/1911.05072)
47. Moussa O, Klakow D, Toneva M. Improving semantic understanding in speech language models via brain-tuning. *ICLR* (2025). [arXiv:2410.09230](https://arxiv.org/abs/2410.09230)
48. Chavarriaga R, Sobolewski A, Millán JdR. Errare machinale est: the use of error-related potentials in brain–machine interfaces. *Frontiers in Neuroscience* 8, 208 (2014). doi:10.3389/fnins.2014.00208

49. Iturrate I et al. Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control. *Scientific Reports* 5, 13893 (2015). doi:10.1038/srep13893
50. O'Sullivan JA et al. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex* 25, 1697–1706 (2015). doi:10.1093/cercor/bht355
51. Ngo H-VV, Martinetz T, Born J, Mölle M. Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron* 78, 545–553 (2013). doi:10.1016/j.neuron.2013.03.006
52. Biasiucci A et al. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. *Nature Communications* 9, 2421 (2018). doi:10.1038/s41467-018-04673-z
53. Cervera MA et al. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. *Annals of Clinical and Translational Neurology* 5, 651–663 (2018). doi:10.1002/acn3.544
54. Cortese S et al. Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J. Am. Acad. Child Adolesc. Psychiatry* 55, 444–455 (2016). doi:10.1016/j.jaac.2016.03.007
55. Dorkenwald S et al. Neuronal wiring diagram of an adult brain. *Nature* 634, 124–138 (2024). doi:10.1038/s41586-024-07558-y
56. Huth AG et al. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458 (2016). doi:10.1038/nature17637
57. Hebart MN et al. THINGS-data, a multimodal collection of large-scale datasets for investigating object representations in human brain and behavior. *eLife* 12, e82580 (2023). doi:10.7554/eLife.82580
58. Obeid I, Picone J. The Temple University Hospital EEG data corpus. *Frontiers in Neuroscience* 10, 196 (2016). doi:10.3389/fnins.2016.00196
59. Deng J et al. ImageNet: a large-scale hierarchical image database. *Proc. CVPR* (2009). doi:10.1109/CVPR.2009.5206848
50. Azevedo FAC et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology* 513, 532–541 (2009). doi:10.1002/cne.21974
51. Sutton R. The bitter lesson. Essay (2019). incompleteideas.net
52. Neurorights Foundation. Safeguarding brain data: assessing the privacy practices of consumer neurotechnology companies. Report (2024). Report (PDF)