



SCROLLY
SCIENCE

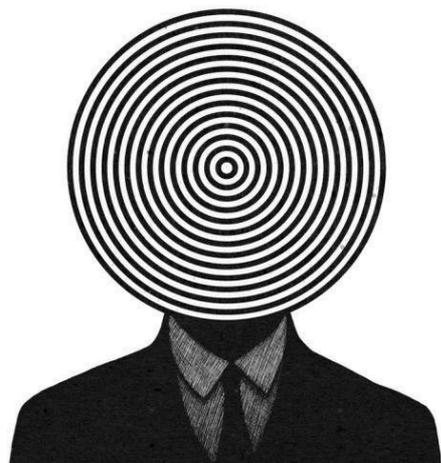
From Brain Signals to Complete Sentences: Captioning the Contents of the Mind

The Challenge of Decoding Thought

For centuries, the contents of our minds have remained fundamentally private—thoughts, memories, and mental images locked away until we choose to express them. But what if that barrier could be crossed?

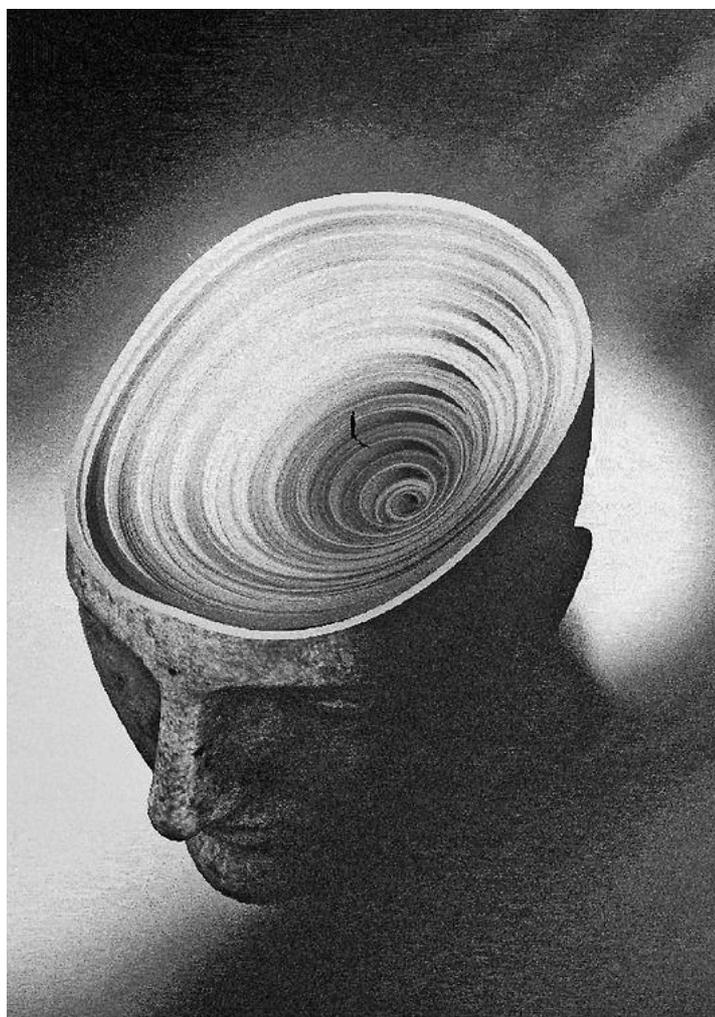
Neuroscientist Dr. Tomoyasu Horikawa has developed a technique called "mind captioning" that translates patterns of brain activity into written sentences. The research demonstrates something remarkable: by combining fMRI scanning with artificial intelligence, meaningful aspects of what we see and remember can be decoded directly from the brain, without speech or language production.

For centuries, the inner workings of the human mind have remained fundamentally private. Thoughts, memories, and mental images form the fabric of our subjective experience, yet they're accessible to others only when we choose to express them through speech, writing, or gesture. But when injury, illness, or disability disrupt these channels, the mind can become isolated from the outside world.



Recent advances in neuroscience and artificial intelligence, however, are beginning to challenge the assumption that internal experience must always remain hidden. In a study published in *Science Advances*, neuroscientist Dr. Tomoyasu Horikawa presents a method he calls "mind captioning" - a technique that translates patterns of brain activity into written descriptions of what a person is seeing or remembering.

The work doesn't claim to read thoughts in the dramatic sense often portrayed in science fiction. Instead, it offers something subtler and arguably more profound: evidence that the meaningful content of mental experience can be inferred from brain activity and rendered into natural language. In doing so, the research opens new avenues for understanding how the brain represents meaning and how those representations might one day support new forms of communication.



Brain decoding isn't a new endeavour. Over the past two decades, researchers have made steady progress in identifying correlations between neural activity and perceptual or cognitive states. Studies have shown that it's possible to determine whether a person is looking at a face or an object, identify basic visual features such as orientation or motion, or even reconstruct rough images from brain scans.

But these achievements have clear limitations. Much of human experience isn't reducible to simple categories or isolated features. We perceive scenes, narratives, intentions, and relationships. We remember past events and imagine future ones, often with rich detail and emotional nuance. Capturing this complexity means shifting away from decoding what is present in the brain towards understanding how meaning itself is represented.

Language provides a natural bridge here. Words and sentences are powerful tools for expressing meaning, but they're not the only way meaning exists in the brain. Long before we speak or write, the brain constructs internal models of the world - integrating sensory input, memory, and expectation into coherent mental representations.

Dr. Horikawa's work is motivated by a central question: can these internal representations be decoded directly, without relying on speech or overt language production? Mind captioning is an attempt to answer that question.

Inspired By Artificial Intelligence

The idea behind mind captioning draws inspiration from a familiar application of artificial intelligence: image and video captioning. Modern AI systems can analyse visual content and generate descriptive sentences such as "a person walking down a busy street" or "a dog playing with a ball in a park". These systems work by extracting semantic features from images and mapping them to language.

Dr. Horikawa reasoned that a similar approach might be applied to the brain. Instead of feeding images into an AI model, they would feed it brain activity recorded while a person experiences those images. The goal wasn't to decode individual words, but to recover the semantic structure underlying the experience.

To test this idea, Dr. Horikawa designed an experiment combining functional magnetic resonance imaging (fMRI) with large language models. fMRI measures changes in blood flow across the brain, providing a detailed picture of which regions are active during specific tasks. While it lacks the temporal precision of other methods, it offers broad spatial coverage - ideal for capturing distributed patterns of activity.

Inside the Experiment

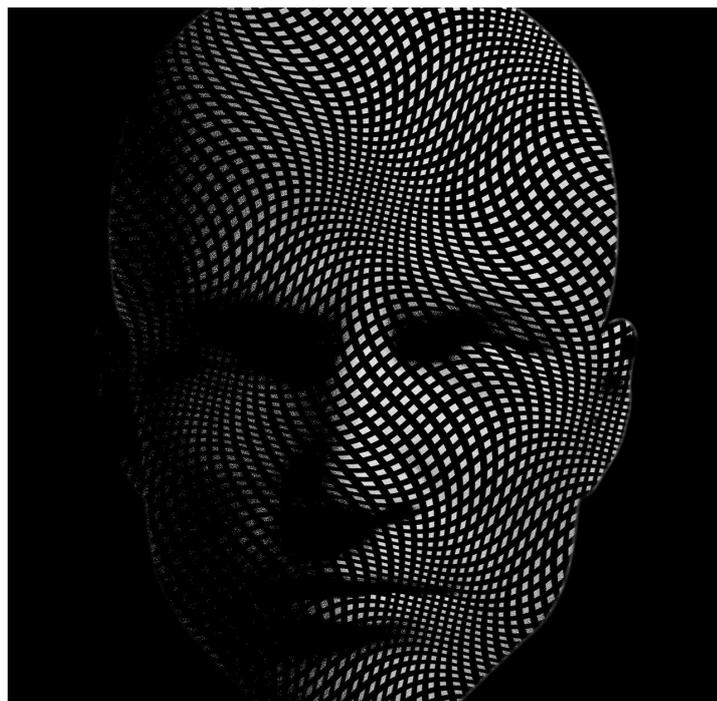
Participants were asked to lie in an fMRI scanner while watching a series of short video clips depicting everyday scenes. The clips showed people moving through environments, interacting with objects, and performing simple actions, and were chosen to reflect naturalistic experiences rather than highly controlled laboratory stimuli.

As participants watched the videos, their brain activity was recorded across the entire cortex. A two-stage computational process was then applied to translate this activity into text.

The first stage involved semantic decoding. Using linear models, Dr. Horikawa learned to map patterns of brain activity onto semantic feature representations derived from a large language model. These features capture high-level aspects of meaning - such as entities, actions, and relationships - typically expressed as sentences or descriptive phrases, rather than isolated words.

The second stage involved text generation and optimisation. The system generated candidate sentences and iteratively refined them so their semantic features aligned as closely as possible with those decoded from the brain data. The final output was a sentence intended to reflect the participant's mental content during the video.

Importantly, this process didn't rely on activity in classical language areas of the brain. Instead, it drew on distributed patterns across visual, associative, and higher-order cortical regions - suggesting that semantic information is represented broadly rather than localised to a single system.



From Perception to Memory

One of the most striking aspects of the study is that mind captioning wasn't limited to immediate perception. In additional experiments, participants recalled previously viewed video clips from memory while undergoing fMRI scanning.

Remarkably, the system could still generate meaningful descriptions of these remembered scenes. This indicates that the decoded signals reflect internal representations of experience, rather than simply mirroring sensory input.

Memory recall is an inherently constructive process. When we remember an event, we don't replay it like a recording; we reconstruct it, drawing on fragments of perception, knowledge, and imagination. The ability of mind captioning to capture aspects of this reconstructed content suggests that semantic representations persist across perception and memory - a finding with significant implications for cognitive neuroscience.

What the Results Reveal

The captions generated by the system weren't perfect replicas of the original scenes. They sometimes generalised or paraphrased, and they didn't capture every detail. Still, they consistently captured the core meaning of the experience: who was involved, what was happening, and how elements were related.

For example, a caption might describe a person walking along a street or interacting with an object, even if it didn't specify exact colours or precise movements. This level of abstraction mirrors how humans themselves describe experiences, prioritising meaning over raw sensory detail.

The findings support the idea that the brain encodes experience in a structured, semantic form that's accessible to computational decoding. They also suggest that large language models - when used carefully - can serve as powerful tools for interpreting neural data, not by imposing language onto the brain, but by acting as a bridge between neural representations and human-readable descriptions.

Why Mind Captioning Matters

The implications of this research extend beyond academic curiosity. One of the most compelling potential applications lies in assistive communication technologies. For individuals who can't speak or type due to paralysis, neurodegenerative disease, or severe injury, current brain-computer interfaces often rely on slow, effortful methods such as selecting letters one by one. A system capable of translating internal semantic representations directly into descriptive language could offer a more natural and efficient alternative.

At the same time, mind captioning provides new insights into how meaning is organised in the brain. By demonstrating that semantic content can be decoded without relying on language-specific regions, the research challenges traditional views of how language and thought are related.

Dr. Horikawa is careful to emphasise the limitations of his work. fMRI isn't suitable for real-time communication, and the system requires extensive individual training data. Ethical considerations - particularly around privacy, consent, and the potential misuse of neural decoding technologies - must also be addressed as the field advances.

A Cautious Path Forward

Mind captioning demonstrates that meaningful aspects of experience leave measurable traces in brain activity, and that these traces can be interpreted in principled ways. What it doesn't do is allow scientists to read minds in any unrestricted sense.

As neuroscience and artificial intelligence continue to converge, approaches like mind captioning offer a glimpse of a future in which neural data can be translated into forms that are meaningful to humans - not as a replacement for communication, but as a supplement when traditional channels fail.

The challenge now lies in refining these techniques, understanding their limits, and ensuring they're developed responsibly. In doing so, researchers like Dr. Horikawa are helping to illuminate one of the most profound questions in science: how the physical activity of the brain gives rise to the rich inner world of the mind.

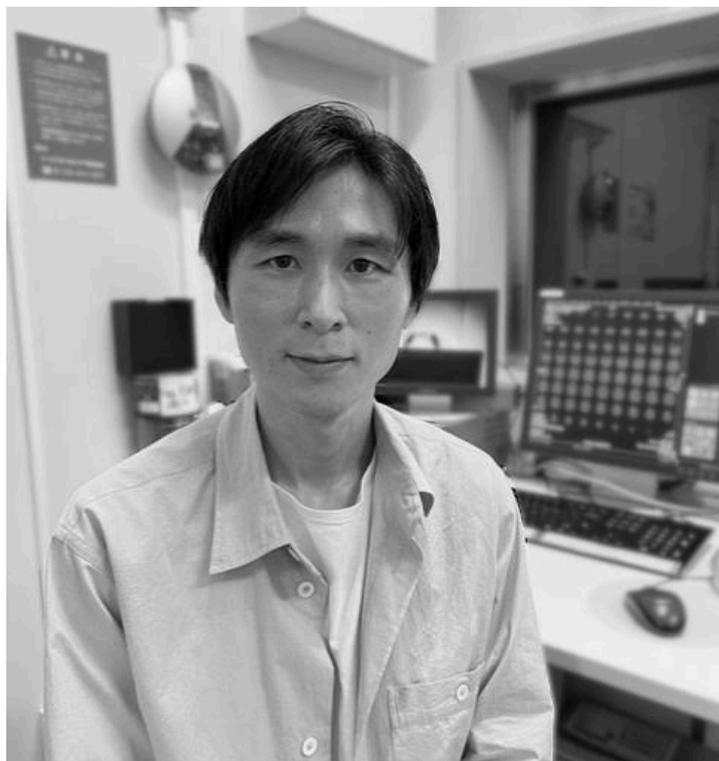


Meet the Researcher

Dr. Tomoyasu Horikawa is a neuroscientist and computational researcher at NTT Communication Science Laboratories in Japan. His work focuses on understanding how the brain represents subjective experience, combining brain imaging techniques with machine learning and computational modelling.

Over the past decade, Dr. Horikawa has been a leading figure in neural decoding research. His studies have explored how visual perception, mental imagery, and memory are encoded in brain activity, often drawing parallels between biological neural systems and artificial neural networks.

By bridging neuroscience and artificial intelligence, his work aims to uncover general principles of representation that apply across both natural and engineered systems.



Research Objectives

At the heart of Dr. Horikawa's research is a fundamental question: how does the brain represent meaning before it's expressed in language? Rather than focusing solely on speech or language production, his work seeks to decode the underlying semantic content of experience itself.

The mind captioning framework reflects this objective. By mapping brain activity onto semantic representations shared with large language models, the approach provides a new way to study cognition, memory, and consciousness while also laying the groundwork for future assistive technologies.

In His Own Words

Your method bypasses the brain's language network entirely, working with non-linguistic mental representations. What was the key insight that led you to this approach rather than trying to decode linguistic information directly?

From the beginning, my goal was to generate accurate, detailed descriptive text of visual content directly from brain activity as a way to investigate how structured, meaning-rich visual semantics are represented in the brain. When considering what would make this work most meaningful, I realized that it would be most impactful if the system could challenge the common assumption that language-like, structured meaning should be represented within the brain's language network itself.

If the system could capture relational and structured semantic information typically expressed in language, without relying on the language network, it would suggest that such meaning is represented more broadly in the brain. This idea was also supported by prior findings showing that individuals with aphasia—who have damage to language-related brain areas—can still perceive and distinguish complex visual events. Taken together, this made the approach both interesting and plausible.

Your participants were native Japanese speakers, yet the system generated accurate English descriptions. What does this suggest about the universality of pre-linguistic mental representations across cultures and languages?

As you point out, a key feature of this approach is that it decodes meaning-based information that is not tied to a specific language. In that sense, the results may suggest that at least some aspects of semantic representation—here, for visual content—are relatively universal and shared across languages.

That said, it remains an open question how strongly these representations are influenced by a person's native language or cultural background. Understanding the extent to which such semantic representations are language-independent versus language-shaped will require more systematic investigation in future work.

The system has only been tested on normal video scenes. What happens when mental content becomes more abstract - for example, emotions, memories, complex reasoning? What breaks down in the current approach?

In principle, a similar approach could be extended to more abstract mental content. For example, memories are already partially addressed in this work through successful decoding of recalled imagery, although this still requires carefully controlled experimental conditions and active participant cooperation.

Emotions and complex reasoning may also be decodable in principle, but doing so would require carefully designed training data. This would include reliable labels reflecting a participant's emotional states or structured tasks capturing different forms of reasoning, along with appropriate textual descriptions. Designing such datasets is non-trivial and needs to be approached carefully, which is why these areas remain future directions.

How would you explain to a general audience the difference between what your system does versus "mind reading" as depicted in science fiction?

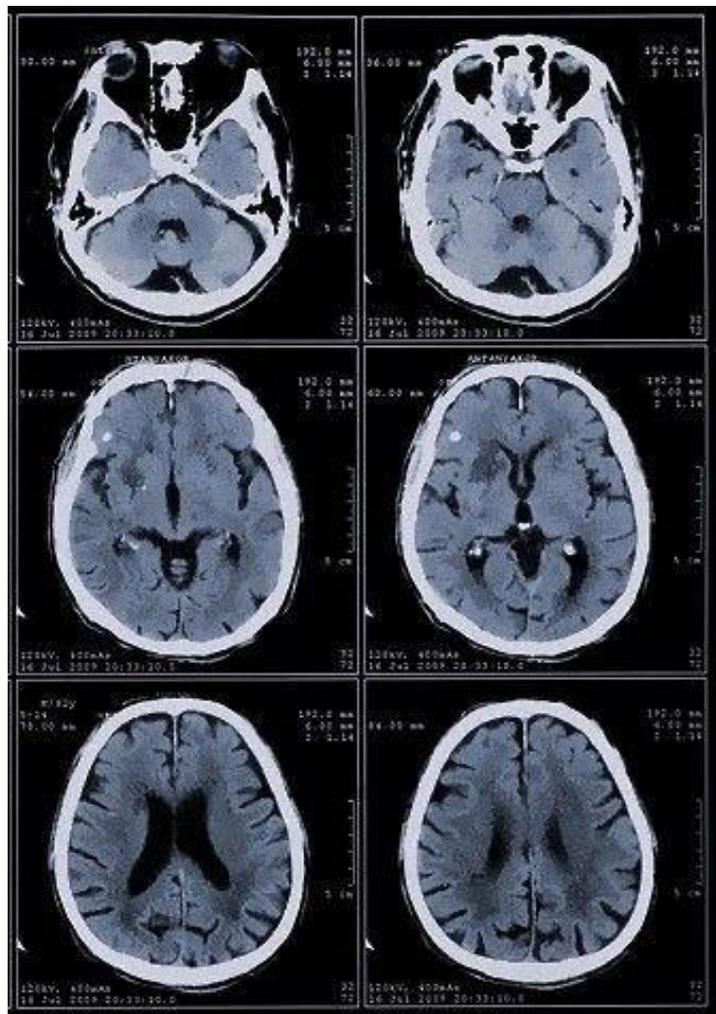
When people hear the term "mind reading," they often imagine a technology that can secretly or involuntarily read someone's thoughts. This is very different from what we demonstrated.

Our approach cannot work without a person's active cooperation. It does not allow someone else to extract thoughts from your brain without consent. Instead, it should be understood as a potential new way for individuals to intentionally communicate their thoughts, under controlled conditions, using brain activity as a signal.

What surprised you most during this research? Was there a moment where the decoded descriptions captured something you didn't expect the system could understand?

What surprised me most was how effective my text optimization approach was when I first tested it. Many previous approaches-including ones I had explored myself-relied on training complex models with large amounts of data to map visual or semantic features to text.

In contrast, the fact that we could achieve strong results without training a new generative model-using word substitution and insertion guided by existing language models-was unexpected. Seeing the method work so well at that early stage was a genuine surprise.



References

Horikawa, T. (2025)
Mind captioning: Evolving descriptive text of mental content from human brain activity. *Science Advances*, 11(3), eadw1464.
<https://doi.org/10.1126/sciadv.adw1464>

Funding

This research was supported by grants from JST PRESTO grant number JPMJPR185B Japan and JSPS KAKENHI grant number JP21H03536.

Competing Interest Statement

Dr. Horikawa declares that he has no competing interests.