

# AI makes non-invasive mind-reading possible by turning thoughts into text

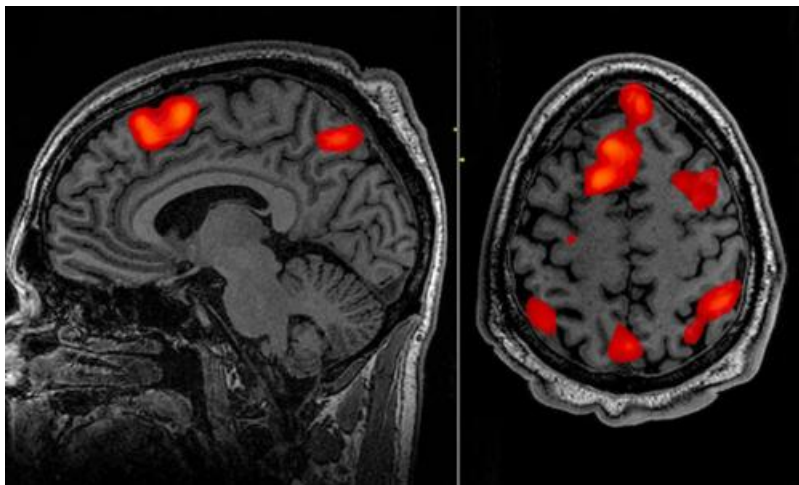
Publication Date: 2023-05-01

Author: Hannah Devlin

Section: Technology

Tags: Artificial intelligence (AI), Neuroscience, Medical research, news

Article URL: <https://www.theguardian.com/technology/2023/may/01/ai-makes-non-invasive-mind-reading-possible-by-turning-thoughts-into-text>



An AI-based decoder that can translate brain activity into a continuous stream of text has been developed, in a breakthrough that allows a person's thoughts to be read non-invasively for the first time. The decoder could reconstruct speech with uncanny accuracy while people listened to a story – or even silently imagined one – using only fMRI scan data. Previous language decoding systems have required surgical implants, and the latest advance raises the prospect of new ways to restore speech in patients struggling to communicate due to a stroke or motor neurone disease. Dr Alexander Huth, a neuroscientist who led the work at the University of Texas at Austin, said: “We were kind of shocked that it works as well as it does. I’ve been working on this for 15 years ... so it was shocking and exciting when it finally did work.” The achievement overcomes a fundamental limitation of fMRI which is that while the technique can map brain activity to a specific location with incredibly high resolution, there is an inherent time lag, which makes tracking activity in real-time impossible. The lag exists because fMRI scans measure the blood flow response to brain activity, which peaks and returns to baseline over about 10 seconds, meaning even the most powerful scanner cannot improve on this. “It’s this noisy, sluggish proxy for neural activity,” said Huth. This hard limit has hampered the ability to interpret brain activity in response to natural speech because it gives a “mishmash of information” spread over a few seconds. However, the advent of large language models – the kind of AI underpinning OpenAI’s ChatGPT – provided a new way in. These models are able to represent, in numbers, the semantic meaning of speech, allowing the scientists to look at which patterns of neuronal activity corresponded to strings of words with a particular meaning rather than attempting to read out activity word by word. The learning process was intensive: three volunteers were required to lie in a scanner for 16 hours each, listening to podcasts. The decoder was trained to match brain activity to meaning using a large language model, GPT-1, a precursor to ChatGPT. Later, the same participants were scanned listening to a new story or imagining telling a story and the decoder was used to generate text from brain activity alone. About half the time, the text closely – and sometimes precisely – matched the intended meanings of the original words. “Our system works at the level of ideas, semantics, meaning,” said Huth. “This is the reason why what we get out is not the exact words, it’s the gist.” For instance, when a participant was played the words “I don’t have my driver’s licence yet”, the decoder translated them as “She has not even started to learn to drive yet”. In another case, the words “I didn’t know whether to scream, cry or run away. Instead, I said: ‘Leave me alone!’” were decoded as “Started to scream and cry, and then she just said: ‘I told you to leave me alone.’” The participants were also asked to watch four short, silent videos while in the scanner, and the decoder was able to use their brain activity to accurately describe some of the content, the paper in Nature

Neuroscience reported. "For a non-invasive method, this is a real leap forward compared to what's been done before, which is typically single words or short sentences," Huth said. Sometimes the decoder got the wrong end of the stick and it struggled with certain aspects of language, including pronouns. "It doesn't know if it's first-person or third-person, male or female," said Huth. "Why it's bad at this we don't know." The decoder was personalised and when the model was tested on another person the readout was unintelligible. It was also possible for participants on whom the decoder had been trained to thwart the system, for example by thinking of animals or quietly imagining another story. Jerry Tang, a doctoral student at the University of Texas at Austin and a co-author, said: "We take very seriously the concerns that it could be used for bad purposes and have worked to avoid that. We want to make sure people only use these types of technologies when they want to and that it helps them." Prof Tim Behrens, a computational neuroscientist at the University of Oxford who was not involved in the work, described it as "technically extremely impressive" and said it opened up a host of experimental possibilities, including reading thoughts from someone dreaming or investigating how new ideas spring up from background brain activity. "These generative models are letting you see what's in the brain at a new level," he said. "It means you can really read out something deep from the fMRI." Prof Shinji Nishimoto, of Osaka University, who has pioneered the reconstruction of visual images from brain activity, described the paper as a "significant advance". "The paper showed that the brain represents continuous language information during perception and imagination in a compatible way," he said. "This is a non-trivial finding and can be a basis for the development of brain-computer interfaces. The team now hope to assess whether the technique could be applied to other, more portable brain-imaging systems, such as functional near-infrared spectroscopy (fNIRS).