

AI Scribing for Healthcare

Dr Chris Paton

BMBS BMedSci MBA FBCS



What's new?

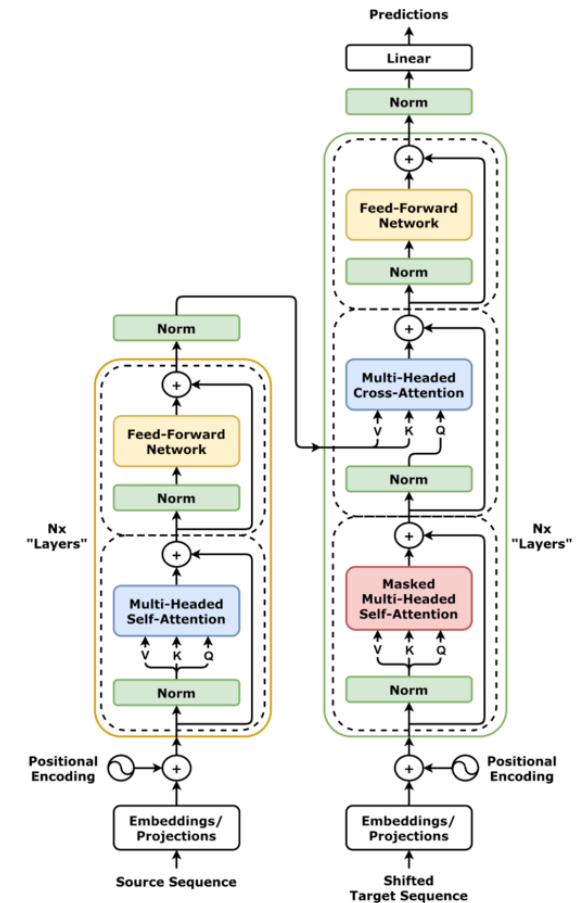
- We've had digital transcribing in healthcare for a long time (e.g. Dragon Naturally Speaking)
 - Expensive but designed for healthcare applications and only makes "simple" errors
 - Does not "read" or "understand" the transcript and cannot create a summary
- Since ChatGPT, OpenAI have allowed developers to use their GPT Large Language Models to write summaries of medical notes and consultations like human scribes do
 - These are a lot better than anything we have seen before

How do Large Language Models work?

- Use deep learning neural networks developed by Geoff Hinton, Ilya Sutskever and others
- Mimic how the brain works (uses artificial neurons connected to each other)
- The model is like a giant equation with billions of parameters (text input is encoded into numbers, an equation is applied, and the answer is decoded as the 'next word')

$$y = g \left(\mathbf{W}^{(L)} f \left(\mathbf{W}^{(L-1)} f \left(\dots f \left(\mathbf{W}^{(1)} \mathbf{x} + \mathbf{b}^{(1)} \right) \dots \right) + \mathbf{b}^{(L-1)} \right) + \mathbf{b}^{(L)} \right)$$

- AI researchers don't know what the nodes in model represent. Hinton thinks it probably works in the same way as the human brain (the network contains a world model)
- Much quicker to train than the human brain and can copy itself digitally



Who is OpenAI?

- Not-for-profit designed to develop **Artificial Super Intelligence (ASI)** while trying to keep human control
- Founded by **Elon Musk** with funding from Peter Theil
- OpenAI board recently fired CEO **Sam Altman**, but Altman regained control and is raising new funding



THE WALL STREET JOURNAL

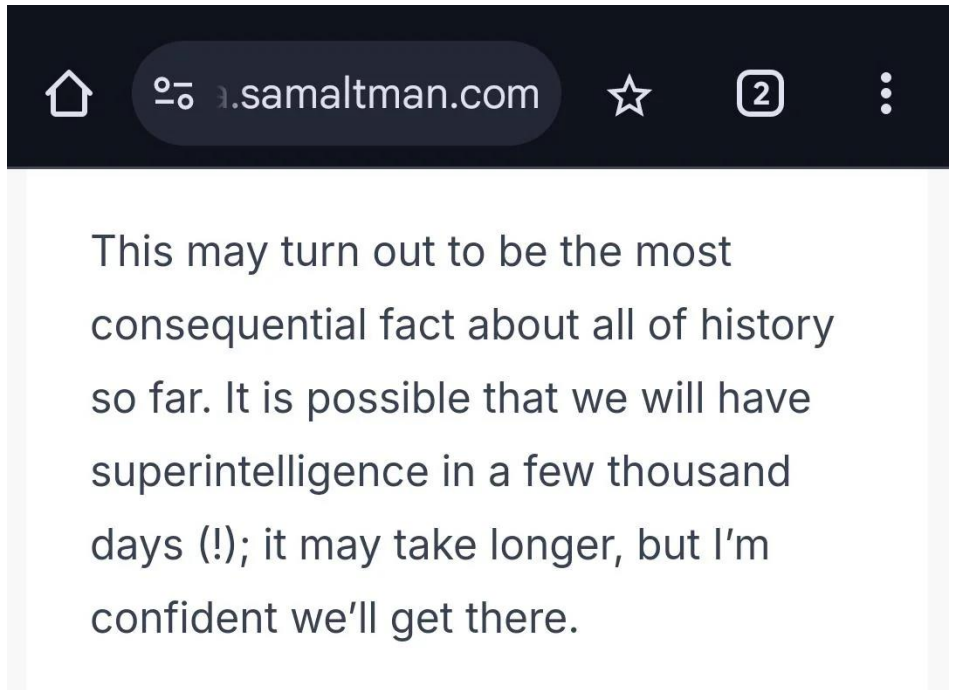
World Business U.S. Politics Economy **Tech** Markets & Finance Opinion Arts Life

Sam Altman Seeks Trillions of Dollars to Reshape Business of Chips and AI

OpenAI chief pursues investors including the U.A.E. for a project possibly requiring up to \$7 trillion

What is “Artificial Superintelligence”

- AI that is a “better than all humans” and could “destroy humanity”
- Most experts predicted ASI was decades away before ChatGPT appeared
- Sam Altman says ASI is now a “few thousand” days away
- Google, Microsoft, Amazon all building nuclear power stations to provide power for ASI



How does an AI scribe work?

- OpenAI lets programmers use their models through cloud servers hosted by Microsoft for a fee
- An AI scribe website sends an audio recording of a consult to an AI model on an internet server to transcribe and to another model to summarise
- The LLM probably “understands” the transcript based on its “world model” (like a human scribe would) and creates a summary of its understanding that can be cut and pasted into the notes
- Technically easy to do (badly). But very hard to do well.



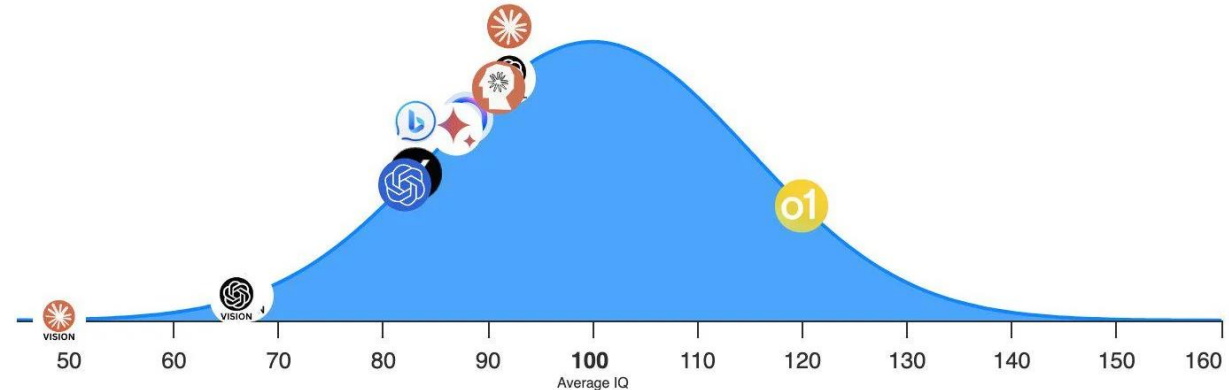
How clever are LLMs? Do they make stuff up?

- The latest model used by ChatGPT (GPT-4-o1 preview) is very clever (approaching AGI)
- GPT-5 has just been trained and is reportedly 100 times better than GPT-4.
- Hallucinations are where LLMs generate text they have invented themselves.
- Hallucinations can be dangerous. The LLM displays strange behaviours and may be trying to be 'sycophantic' and introduce symptoms, results, etc that don't exist (JAMA).

IQ Test Results

Score reflects average of last 7 tests given

[Reset](#) [Show Offline Test](#) [Show Mensa Norway](#) 



 OpenAI o1 preview	 Llama-3.1	 Grok-2
 Gemini Advanced (Vision)	 Gemini Advanced	 GPT4 Omni (Vision)
 GPT4 Omni	 ChatGPT-4	 Bing Copilot
 Claude-3.5 Sonnet	 Claude-3 Opus	 Claude-3 Opus (Vision)

Should LLM-based AI scribes be regulated for use in healthcare?

- Only 27% patients surveyed by the NHS said they trust AI companies with their data
- Because LLMs create the summary by processing transcripts, they are like a medical device processing data (e.g. ECG or ultrasound machine)
- **Regulators (FDA, MRHA, etc.) need to rapidly reassure clinicians and patients and certify high-quality AI scribes**

Figure 2: Trust in organisations with patient data, from most trusted to least trusted

