



Born Supremacy – AI as a Pale Shadow of Real Humanity?

Professor Matt Jones

17 March 2026

Artificial intelligence systems are often presented to us as rivals. We are told they diagnose diseases more accurately than doctors, write essays faster than students, analyse data more efficiently than researchers, and may soon outperform humans across many intellectual domains. Across this lecture series so far, we have explored the unsettling futures that such claims suggest. We have asked whether AI might become an overlord, whether we might eventually merge with it, or whether we might gradually become domesticated by the conveniences it offers.

In this fourth lecture, I want to reframe that story. The extraordinary achievements of modern AI are real and should not be dismissed but the way we interpret them matters. Much of what currently impresses us in artificial intelligence is a matter of performance—producing convincing outputs quickly and accurately. Human intelligence, however, is something richer and more complex. It is not only performance but inhabited intelligence: intelligence embodied in living beings, unfolding through memory, emotion, responsibility, and shared practices over time.

To explore this idea, I began with an image from popular culture. At the beginning of *The Bourne Identity*, the character Jason Bourne is pulled from the sea, apparently dead but soon revealed to possess astonishing capabilities. He can fight, evade enemies, analyse threats and respond with extraordinary precision. Yet he does not know who he is. He has skill without memory, performance without identity. The story of the films is not simply about defeating adversaries but about rediscovering continuity: understanding who he has been and what his life means.

This metaphor captures something important about many contemporary AI systems. They often display seductive competence, but competence alone is not the same as intelligence as human beings live it. Humans do not simply produce correct outputs. Our intelligence unfolds through bodies, histories, emotions, and relationships.

Consider the achievements of figures such as Marie Curie, Yo-Yo Ma, or Lionel Messi. Their intelligence cannot be reduced to solving isolated tasks. Curie's discoveries emerged from years of experimental persistence. Yo-Yo Ma's artistry reflects decades of physical practice intertwined with emotional interpretation. Messi's brilliance on the football field depends on instantaneous perception, bodily skill, strategic anticipation, and deep experience of the game. Their achievements illustrate that human intelligence is enacted through practice—through the integration of mind, body, and world.

Yet the same principle applies more quietly to all of us. Human intelligence is never purely abstract. Our brains, our senses, our bodies, and our emotions continually shape how we understand and respond to the world around us.

Modern neuroscience helps illuminate this. For a long time, perception was imagined as if the brain were looking out through a clear window onto the world. Contemporary theories suggest something more dynamic. For instance, neuroscientist Karl Friston has detailed how the brain continually predicts what it expects to encounter and then adjusts those predictions in response to incoming sensory signals. In that sense, perception is not passive recording but active modelling. Our brains generate expectations and then refine them through interaction with reality.

This idea has been summarised provocatively as the claim that perception is a kind of “controlled hallucination.” The brain generates a model of the world, but the real world constantly corrects and constrains that model.

Interestingly, this predictive process bears a superficial resemblance to the way generative AI systems operate. Large language models are trained by attempting to predict missing or next words in vast quantities of text. When their prediction differs from the actual word, the model adjusts its internal parameters to reduce error. Through extensive and iterated training, the system becomes increasingly good at predicting plausible continuations of language.

However, human predictive processing is embedded in a living organism whose survival and well-being depend on accurate perception. We predict the world not merely to produce correct statements but to navigate dangers, pursue goals, and sustain relationships. Our predictions matter because we inhabit the world and are vulnerable within it. So, Jason Bourne in the books and movies is continually seen trying to anticipate the outcome of tense and fast-paced outcomes, where bad predictions might end in his death.

Language provides another comparison. Human listeners, like language models, often anticipate the next word in a sentence before it arrives. Yet when we predict language, we frequently do more than calculate probabilities. We imagine, simulate, and even feel. If someone describes slicing a tomato, many listeners inwardly experience the visual texture, the resistance of the knife, the smell of the fruit. If someone describes walking into an exam room and suddenly realising a pen has been forgotten, many people feel a flicker of anxiety. Our understanding of language is entangled with memory, perception, and emotion.

Memory itself provides examples of further differences. Human working memory—the information we can actively hold while performing a task—is surprisingly limited. Psychological research suggests we can manage only a small number of “chunks” at once. Yet each chunk can unlock vast networks of meaning. A short phrase, a melody, or a familiar acronym may evoke a complex world of associations and experiences.

Large language models have an analogous mechanism known as a **context window**, the amount of text they can consider while generating output. In modern systems this window may span the equivalent of hundreds of pages of a book. At first glance this appears to dwarf human working memory. But the comparison is misleading. Human memory chunks are deeply connected to lived experience, emotion, and personal history in ways that simple word sequences cannot capture.

Human semantic memory—our knowledge of concepts and facts—forms networks of associations. Language models similarly represent words as vectors within high-dimensional mathematical spaces where related meanings cluster together during the training process. These structures allow models to capture patterns of language effectively. Human concepts,

though, are not merely abstract patterns. They are grounded in perception, action, culture, and experience.

Procedural memory affords another perspective on the distinction inhabitation gives us in comparison with frontier AI. When humans learn to cycle, play an instrument, or perform surgery, the skill becomes embodied rather than embedded as a statical pattern. It transforms not only the efficiency of movement but the person's sense of identity and capability. Machine learning systems can acquire motor skills through reinforcement learning, and other forms of algorithm, but the meaning of mastery—the pride, the struggle, the transformation of self—remains distinctively human.

Perhaps the most clear differences appear in episodic memory, our capacity to remember personal experiences. Human memory is not a simple replay of stored events. Psychological research, including classic experiments by Elizabeth Loftus, shows that recollection is reconstructive. Each act of remembering involves rebuilding a narrative from fragments of the past. Yet these memories matter deeply because they form the story of our lives—our identities, relationships, and responsibilities.

Human intelligence is also fundamentally social. It emerges through collective practices such as music, sport, science, and politics. These practices involve shared norms, trust, accountability, and stakes. A scientific community, for example, is not merely a collection of correct statements but a network of people who test, challenge, and refine knowledge together. AI systems may process enormous amounts of scientific literature, but processing information is not the same as participating in a community of inquiry.

This distinctions we explore in the lecture are important provocations when considering the future of AI. The central question is not simply how capable machines can become. It is also what kind of relationship we wish to cultivate between human and artificial intelligence. If our goal is merely to maximise machine capability, we risk building systems that erode human judgment, responsibility, and expertise. A more constructive vision is to design AI that supports human agency—tools that sharpen our thinking, extend our capabilities, and keep us accountable for the decisions we make.

Near the end of the lecture, I reflected on the work of Desmond Tutu and South Africa's Truth and Reconciliation Commission. One could imagine feeding vast quantities of data about conflict and violence into an optimisation system to generate policies that minimise unrest. Yet what the country required was something different: moral repair. That process required people to gather, to speak, to remember, to weep, and to forgive. It required human presence and shared vulnerability. This reminds us that intelligence is not only about producing correct answers. It is about living meaningfully in a shared world.

Returning to Jason Bourne, the story becomes satisfying not because of his technical skill but because he ultimately recovers his identity. The analogy invites us to reflect on today's AI systems. They may demonstrate impressive performance, but performance alone does not constitute the full richness of intelligence.

Human beings remain inhabited intelligences—people whose thinking is inseparable from memory, embodiment, emotion, and community. If intelligence were merely output, then the story might already be over. But intelligence, as we have traced it, is enacted in temporally extended, norm-governed practices. It requires agents who remember, anticipate, and can be held accountable.

So when we ask, "If machines can do these things, what are we for?" the answer is not defensive. In this present moment when there is a blazing hot "AI Summer" we instinctively feel that we are very much in the shade; but it is AI and not us which is the "pale shadow".

This insight may help us remember what we are for: not merely to optimise outcomes, but to create meaning, exercise judgment, and shape the future together.

So, this lecture is not anti-AI or anti-automation. It is pro-agency. It recognises that intelligence enacted in time and under consequence – human intelligence - is the essential medium through which transformative creativity, innovation and purposes can emerge. If we recognise and assert this, we will design powerful AI tools that lead to surprising, delightful, transformative people, communities and societies.

© Professor Matt Jones 2026

References and Further Reading

For those who would like to explore the ideas behind this lecture in more depth, the following works provide accessible entry points into the science and philosophy underpinning the argument.

Perception as Prediction

- Rao, R. P. N., & Ballard, D. H. (1999). *Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects*. *Nature Neuroscience*, 2(1), 79–87. <https://doi.org/10.1038/4580>
- Friston, K. (2010). *The free-energy principle: a unified brain theory?* *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Clark, A. (2013). *Whatever next? Predictive brains, situated agents, and the future of cognitive science*. *Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>

Memory and Future Simulation

- Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). *Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration*. *Neuropsychologia*, 45(7), 1363–1377. <https://doi.org/10.1016/j.neuropsychologia.2006.10.016>
- Schacter, D. L. (2001). *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Boston: Houghton Mifflin.
- Loftus, E. F. (1979). *Eyewitness Testimony*. Cambridge, MA: Harvard University Press. (Reissued with new preface, 1996.)

Embodied and Situated Cognition

- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Collective Intelligence and Norms

- Merton, R. K. (1973). *The normative structure of science*. In N. W. Storer (Ed.), *The Sociology of Science: Theoretical and Empirical Investigations* (pp. 267–278). Chicago: University of Chicago Press. (Original work published 1942)
- Oreskes, N. (2019). *Why Trust Science?* Princeton, NJ: Princeton University Press.

AI Systems and Large Language Models

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. *Nature*, 521, 436–444.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. In *Advances in Neural Information Processing Systems 30*, 5998–6008.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). *Language models are few-shot learners*. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 1877–1901. arXiv:2005.14165

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, pp. 610–623. <https://doi.org/10.1145/3442188.3445922>

Design and Human Agency in AI

Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford: Oxford University Press.

Bridle, J. (2018). *New Dark Age: Technology and the End of the Future*. London: Verso.

Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking/Penguin.