

WHY AI HYSTERIA MISSES HUMANITY'S TRUE POTENTIAL

June 2024



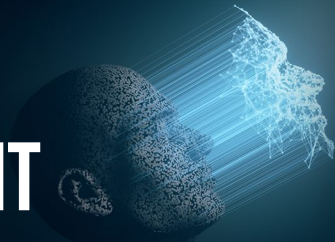
CONTRIBUTORS

David Brunner

CEO & Founder, ModuleQ

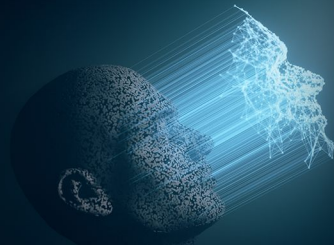
Anupriya Ankolekar PhD

Co-Founder & Principal Scientist, ModuleQ



Content

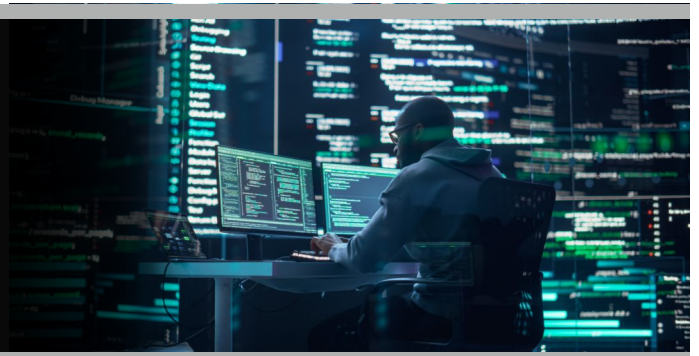
Introduction	3
AI's Illusion of Omniscience	4
AI's Limited Superpowers	5
Knowledge Creation	6
AI-Human Collaboration	7
About the Authors	9
References	10



Across the large western economies, citizens are broadly opposed to AI. This is the troubling conclusion of the 2024 Edelman Trust Barometer. In the US, the UK, Germany, and France, less than 20% of the population “embrace the growing use of AI”, while 50% or more “reject the growing use of AI.” Given the tremendous and transformative potential of AI, this widespread antagonism is alarming. Unchecked, it may threaten the prosperity and sustainability of these societies. The Barometer finds many underlying concerns. At the top of the list are threats to privacy and “the possibility that it may devalue what it means to be human”.

While privacy has been the subject of much discussion and legislation, devaluing “what it means to be human” seems to be a thornier challenge. Extrapolating from the extraordinary achievements of Large Language Models (LLMs) such as GPT-4, prominent scientists and entrepreneurs speculate that AI will surpass human intellect, rendering human beings superfluous and obsolete. To the extent that our sense of self-worth comes from our ability to make unique and valuable contributions to society, that would certainly devalue “what it means to be human”.

The unique abilities of human beings are far more profound, powerful, and mysterious than generally acknowledged



But this fear is unfounded, a misguided conclusion drawn from two mistaken premises: first, an overestimation of the abilities of AI and, second—equally important but often overlooked—an underestimation of humanity. The abilities of AI are extraordinary and superhuman, yet they are also much narrower than they appear. Conversely, the unique abilities of human beings are far more profound, powerful, and mysterious than generally acknowledged. This is likely because of the great influence of economic theory in public discourse, which is premised on assumptions about human behavior that obscure the complexities at the heart of our humanity.

Is the reductive caricature of human nature at the core of economic science why citizens in strongly neoliberal nations such as the US, the UK, and Germany are much more hostile toward AI than those of Japan, India, and China?

Perhaps a more nuanced view of AI’s limitations and humanity’s self-actualising power may help overcome the fear of AI and facilitate broader adoption of AI in ways that complement human work and serve the broader interests of humanity.



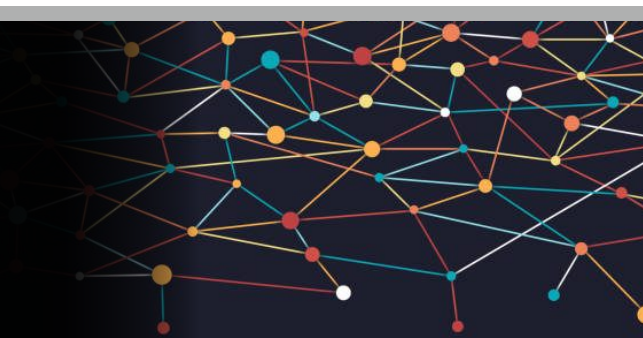
AI's illusion of omniscience

AI, in the form of LLMs, has achieved superhuman mastery of certain quintessentially human skills: verbal composition, persuasion, and rhetoric. Trained on a larger volume of written language than any human could ever hope to absorb and adapted to human preferences by “reinforcement learning from human feedback” (RLHF), LLMs demonstrate an awe-inspiring mastery of language. The most advanced models answer nearly any question, glibly and persuasively, in any major language, with impressive eloquence. LLMs can perform wondrously clever tricks of verbal dexterity, effortlessly revising their answers into rhyming sonnets and limericks, or even adept imitations of famous authors from Shakespeare to James Joyce. Essays and explanations that would take humans many hours of painstaking effort to compose appear on the screen in seconds.

This extreme verbal proficiency gives the impression of near omniscience, and it exceeds human ability, at least in the sense of rhetorical impact. In a recent experiment², researchers invited human subjects to debate controversial topics. The subjects were assigned at random to debate, via text messaging, either another human or an AI. The most effective AI turned out to be far more persuasive than humans. Interestingly, the AI also argued more logically and analytically.

Is it any surprise, then, that when humans encounter LLMs, they wonder if they are face-to-screen with a new class of being that will consign human beings to the scrap heap of evolution? Researchers find that **when humans collaborate with LLMs, they struggle with feelings of inferiority** and diminished ownership of their work³. In the words of one study subject, “if the machine is as good as me, then what use am I?”.

**When humans collaborate
with LLMs, they struggle with
feelings of inferiority**



The reality of AI's limited superpowers

Yet, the superhuman abilities of AI are narrower than they appear. The tendency of LLMs to “hallucinate” has been widely remarked. The term is not entirely apt: LLMs are masters of

rhetoric, of creating patterns of words that are pleasing and compelling. But they create these patterns based only on statistical likelihood, without any subjective intent or regard for factual knowledge or morality. LLMs resemble the master rhetorician Gorgias, whose dialogue with Socrates is retold by Plato. A practitioner of rhetoric, says Socrates, “never has to know the actual facts of any issue; instead, he’s equipped himself with a persuasive ploy which enables him to make non-experts believe that he knows more than experts”.⁴

As research probes the limitations of LLMs, the aura of omniscience is fading. In a study of ChatGPT usage at the Boston Consulting Group⁵, researchers tasked management consultants with solving a business problem. The consultants were given interview transcripts and quantitative data. The quantitative data had been manipulated to appear complete, while omitting crucial details. It is only with careful critical analysis of the transcripts that consultants spot the omissions and reach the correct answer.

Consultants using ChatGPT completed the task faster, and their answers were judged by human reviewers to be of higher “quality,” but they were more likely to fall into the analytical trap: They got the correct answer only 65.5% of the time, compared to 84.5% for consultants without ChatGPT. **LLMs may appear to articulate themselves logically and analytically, but they lack critical thinking skills.**

At ModuleQ, our research team found that **LLMs performed very poorly at analytical questions that would be easy for junior financial analysts**⁶. For example, LLMs could not reason accurately about the calendar quarter in which financial transactions occurred, even when given a table containing the closing dates of the transactions. Although it was possible to improve LLM performance by extending them with specialised analytical tools, our findings highlight the fallibility of LLM “reasoning”.

With **AI proving to be more superficial rhetorician than omniscient genius**, the unique qualities of human beings tend to be underestimated, especially in the economic sphere. And while economic productivity is only a narrow slice of “what it means to be human”, we focus here on the world of knowledge-intensive work, since that is where AI is generally expected to be particularly disruptive.

**AI proving to be more
superficial rhetorician
than omniscient genius**



The upper hand: Knowledge creation and human subjectivity

There is an overriding tendency in economic science to ignore most of “what it means to be human” and reduce human behavior to maximisation of profits (in the case of business organizations) and self-interest (in the case of individuals). Yet, as leading organisational theorist Ikujiro Nonaka points out, **human subjectivity is fundamental to how organisations create the knowledge that shapes their daily activities and their destinies.**

If organisations were nothing more than the profit-maximising machines of economic theory, perhaps future generations of AI, augmented with knowledge bases and reasoning engines, could someday replace humans. Nonaka and Toyama’s “Theory of the Knowledge-Creating Firm” explains why this will not happen.⁷

Organisations are communities of people engaged in complex processes of knowledge creation, driven by ideals, interactions, shared experiences, and social capital. Through the SECI process of socialisation, externalisation, combination, and internalisation (Figure 1), humans acquire tacit knowledge. This tacit knowledge becomes encoded as explicit knowledge assets, and explicit knowledge catalyses more generation of tacit knowledge. This process is inextricably grounded in human subjectivity, intent, values, and interactions.

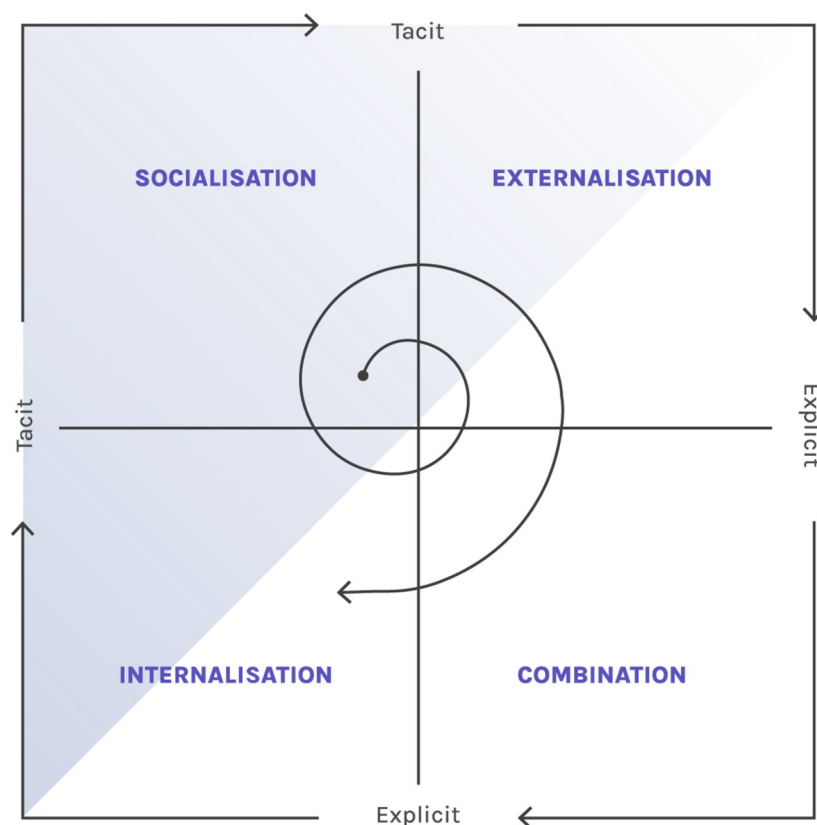


Figure 1: The SECI model of knowledge creation (Source: ModuleQ, based on research by Ikujiro Nonaka and Hirotaka Takeuchi)

In Nonaka and Toyama's words:



Since knowledge emerges out of subjective views of the world, it probably cannot be reached by the one and only absolute 'truth'. The knowledge-creating process is idealistic, since knowledge is created through the social justification process, which relentlessly pursues a truth that may never be reached. We can say that the theory of knowledge creation is based on an idealistic pragmatism which synthesises the rational pursuit of appropriate ends, whose appropriateness is determined by ideals.

Knowledge creation is not limited to senior executive leadership teams or research labs, according to Nonaka and Toyama. Organisational knowledge creation is the essential activity of all growing businesses, whether they are software developers, automobile manufacturers, or retail stores. It is a uniquely human activity that AI cannot in any way replace, because it is social, idealistic, and embodied in human experience.

AI-human collaboration

If “what it means to be human” transcends rhetoric and reasoning, is AI merely a tool for more process automation, or could it enhance our uniquely human processes of organisational knowledge creation?

Early research suggests that AI working alongside humans, termed “co-intelligence” by applied generative AI expert Ethan Mollick⁸, could enhance our lives as human workers. Economist and human-centered AI expert Erik Brynjolfsson and his colleagues investigated the use of generative AI in assisting customer service agents.

The AI provided human customer service agents with suggested responses and potentially relevant technical documentation⁹. The human agents were free to use the AI-generated suggestions or not, at their own discretion. The study found that the AI system increased agent productivity by 13.8%. Perhaps more interestingly, the AI system appeared to help disseminate knowledge among the agents and help newer agents internalise knowledge more rapidly, enhancing the performance of the SECI cycle. In addition, the AI appeared to help agents get along better with customers, improving the sentiment of interactions, reducing agent turnover, and reducing customer frustration, as evidenced by fewer demands to escalate conversations to supervisors.

Customer support is a relatively routine job, but AI may also impact knowledge dissemination and learning in more open-ended contexts. In a field experiment with 640 Kenyan

AI support is only as valuable as the knowledge that it curates and diffuses



entrepreneurs, researchers studied the effects of giving entrepreneurs access to an AI business advisor¹⁰. Interestingly, this improved the business performance of high-performing entrepreneurs by 15%, but decreased that of low-performers by 8%.

The researchers report that low-performers sought advice about more difficult business problems (probably because their businesses were struggling), while the high-performers asked more questions about how to grow and expand (likely because their businesses were performing well).

This suggests that **AI support is only as valuable as the knowledge that it curates and diffuses**. Notably, in the Brynjolfsson et al. study, the AI suggestions were informed by the outcomes of past support interactions. In contrast, the AI business advisor used an LLM with domain-tuned prompting to answer open-ended questions from the study participants.

Taken together, recent research on the strengths and weaknesses of LLMs and AI, in combination with theories of the complex, subjective, and social processes through which organisations create knowledge, provide a starting point for a more balanced and grounded discussion of AI that neither overestimates AI nor underestimates humans.

There are legitimate concerns about the rhetorical brilliance of LLMs, concerns that would have been familiar to Socrates over two millennia ago. There are also concerns about AI diffusing ineffective or inaccurate “knowledge” that may be harmful to individuals and society. Nevertheless, applied judiciously, AI can augment humans in our working lives, helping us learn faster and build higher-functioning organisations that effectively balance economic objectives with more humanistic ideals.

ABOUT THE AUTHORS



David Brunner
CEO & Founder
ModuleQ

I write about AI, innovation, and organisational learning.



Anupriya Ankolekar PhD
Co-Founder &
Principal Scientist
ModuleQ

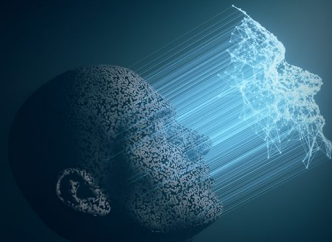
I work on human-centred AI, knowledge-based systems and human-computer interaction.

The views expressed here are their own and do not necessarily reflect the views of company or its staff.

To access more reports, scan the QR code below:



REFERENCES



- ¹ Edelman Trust Institute. “2024 Edelman Trust Barometer Supplemental Report: Insights for the Tech Sector.” 2024.
- ² Francesco Salvi et al. “On the Conversational Persuasiveness of Large Language Models: A Randomized Controlled Trial.” arXiv, 2024.
- ³ Charlotte Kobiella et al. “‘If the Machine Is As Good As Me, Then What Use Am I?’ -- How the Use of ChatGPT Changes Young Professionals’ Perception of Productivity and Accomplishment.” arXiv, 2024.
- ⁴ Plato. Gorgias, 459c. Translated by Robin Waterfield. Oxford University Press, 1994.
- ⁵ Fabrizio Dell’Acqua et al. “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality.” Harvard Business School Working Paper 24 -013, 2024.
- ⁶ Nicholas Harvel et al. “Can LLMs Answer Investment Banking Questions? Using Domain-Tuned Functions to Improve LLM Performance on Knowledge-Intensive Analytical Tasks.” Forthcoming in Proceedings of the AAAI Spring Symposium on Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge, Stanford University, California, USA, March 25 -27, 2024.
- ⁷ Ikujiro Nonaka and Ryoko Toyama. “The theory of the knowledge -creating firm: subjectivity, objectivity
- ⁸ and synthesis.” Industrial and Corporate Change, Vol 14, No 3, pp. 419–436, 2005
- ⁹ Ethan Mollick. Co-Intelligence: Living and Working with AI. Penguin Random House, 2024.
- ¹⁰ Erik Brynjolfsson et al. “Generative AI at Work.” National Bureau of Economic Research working paper 31161, 202 3.
- ¹¹ Nicholas Otis et al. “The Uneven Impact of Generative AI on Entrepreneurial Performance.” OSF Preprints, 2024.