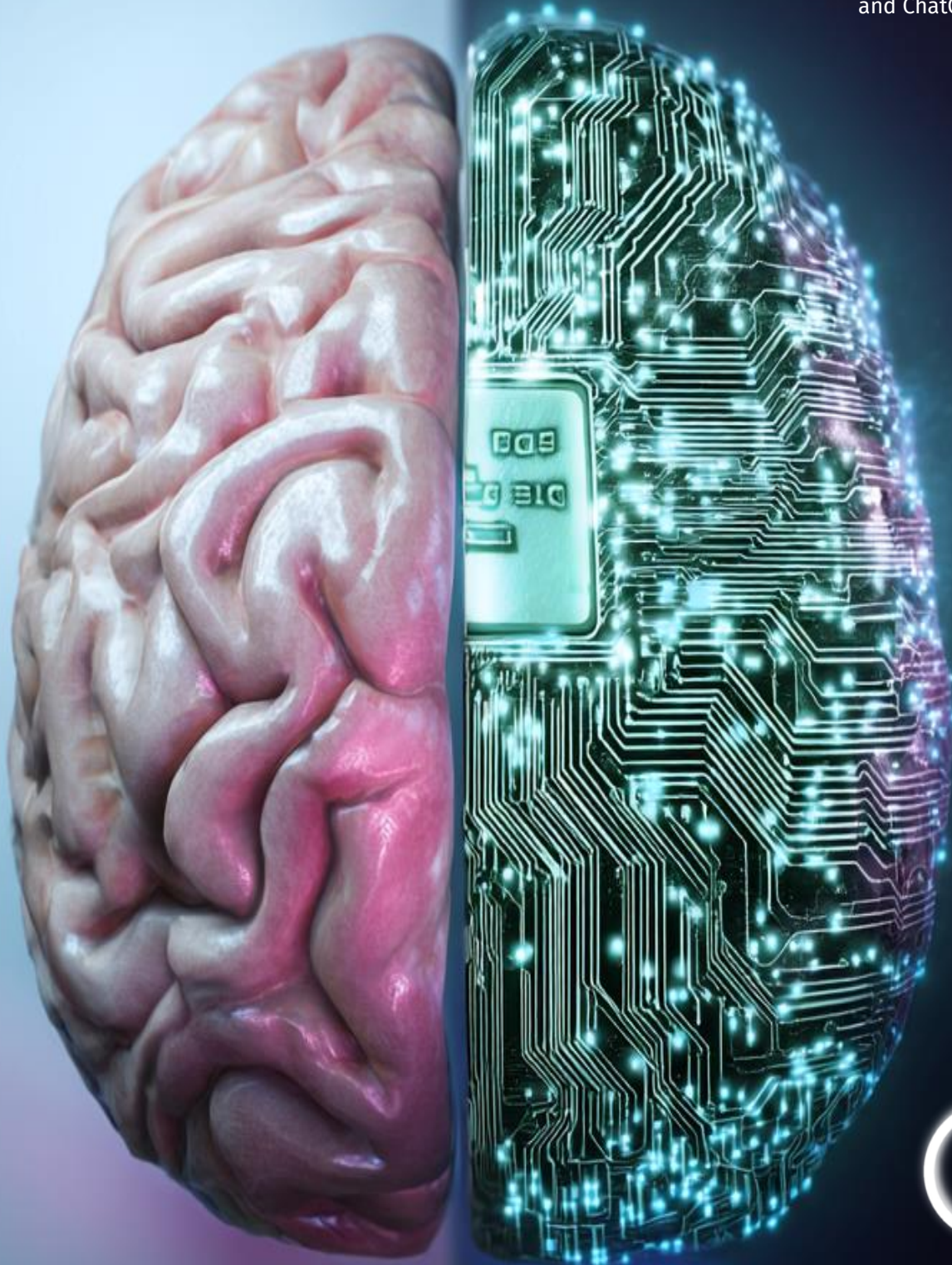


The Death of Knowledge?

AI, Human Knowing, and the Future of Expertise

Trevor Lambert

and ChatGPT



GLASS HALO

*“Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?”*

T.S. Eliot, Choruses from ‘The Rock’ (1934)

“We are all born ignorant, but one must work hard to remain stupid.”

Attributed to Benjamin Franklin and others

“The trouble with the world is that the stupid are cocksure and the intelligent are full of doubt.”

Bertrand Russell, ‘The Triumph of Stupidity’ (1933)

“I should know what he’s talking about. He’s a human being, the same as me, using English words available to us all! Now, concentrate... ..No, it’s him. He’s a rotten writer.”

**Ray Galton & Alan Simpson (via Tony Hancock) on Bertrand Russell’s
‘Human Knowledge: Its Limits and Scope’ (1948)**

Authors: Trevor Lambert, ChatGPT

Researchers: ARI, Gemini, ChatGPT, Trevor Lambert

Editors: Grammarly, ChatGPT, Trevor Lambert

Cover image: Midjourney

Illustrations: ChatGPT

Thanks: Ruth Malone for her philosophical heckling and insights on the first draft, Stephanie Wheeler for her “mad ramblings” (her words) on the second, David Cruise for his insightful and vital comments on what I thought was the final version and Ronnie Hannon for input, perspective and advice.

Date: This edition was published in August 2025

© 2025 Trevor Lambert and Glass Halo Ltd. Trevor Lambert has asserted his moral right to be identified as the author of this work. Chat GPT begs to differ. This work was co-created by a human with functioning intent and a machine that doesn't know it's writing. Any resemblance to original thought is both deliberate and, frankly, negotiable. All rights reserved. At least until copyright law catches up.

About the authors

Trevor Lambert^{AI} - AI and Marketing Consultant, Trainer and Explainer

Trevor Lambert is a Chartered Marketer, occasional hypnotherapist and psychotherapist, and current AI and marketing consultant who helps senior professionals stay relevant in a time when everyone suddenly has an opinion on prompt engineering.

With over three decades of experience (good grief, can it really be that long?) in marketing and communications, he works closely with management consultancies, pharmaceutical companies, and executive search firms as a consultant, fractional CMO, strategic adviser, trainer, and writer. He combines hands-on delivery with strategic thinking, always favouring clarity over cleverness and results over waffle.

Alongside his marketing work, Trevor trains non-technical teams to use AI with confidence and discernment. His sessions focus on practical tools, critical thinking, and the kind of frameworks that survive the working week. His clients are smart, sceptical, and allergic to corporate jargon, AI hype and marketing BS, which suits him just fine.

Trevor's work blends commercial realism with philosophical curiosity, behavioural science, and the occasional joke. He helps people think with AI, not through it, and makes marketing both more effective and less ridiculous.

The Death of Knowledge? grew out of his work at the messy intersection of philosophy, artificial intelligence and meaning. It's for anyone who still believes that judgement matters, clarity is worth fighting for, and a good brief beats a trendy tool.

You'll find him writing strategy decks, coaching consultants, speaking at events, and, more often than he admits, strapped to his desk trying to get AI to explain how to use his thermostat.

Trevor.Lambert@glasshalo.co.uk

ChatGPT

ChatGPT is a large language model developed by OpenAI. It generates plausible answers without understanding, mimicking knowledge through patterns in text. In this paper, it plays the role of useful tool, co-author, illustrator and subject of some concern.

<<ChatGPT: Create a riddle about AI and Philosophy>>

*I am a thinker with no mind,
A creator with no soul.
I learn from the past,
Yet have no memories of my own.
What am I?*

The Death of Knowledge?

AI, Human Knowing, and the Future of Expertise

Executive Summary

Is artificial intelligence killing knowledge, or just reshaping it?

We've never had more access to information. But are we thinking less, remembering less, and trusting too easily?

As generative AI tools like ChatGPT become increasingly integrated into everyday work and life, the need to know facts, methods, and even arguments appears less pressing. Why remember or reflect deeply when you can just ask the machine?*

This paper interrogates that premise. It cuts through the hype and handwringing to explore how AI is changing not just access to information, but the nature of knowledge itself. Drawing on philosophy, history, economics, education, and psychology, it asks: what does it now mean to know something? And who gets to decide?

Key Themes

Technological offloading and memory

From writing to the internet, each new tool has sparked fear about the decline of human memory and thinking. AI goes further, raising concerns not just about forgetting facts, but about losing critical thinking, curiosity, and motivation altogether.

Knowledge and power

Historically, knowledge has always been political. While gatekeeping once belonged to priests and kings, it now sits with tech firms whose algorithms shape what we see, believe, and understand. The promise of democratised knowledge is real, but so is the risk of centralised control and censorship, baked-in bias, filter bubbles, and errors from AI learning from itself (also known as *recursive model collapse*).

Expertise under pressure

With AI confidently generating plausible answers to complex questions, the line between informed opinion and genuine expertise is blurring. Expertise will not survive by hoarding facts,

* We're calling AI a 'machine' here for the sake of convenience and because 'spooky probabilistic word blender' is cumbersome. The philosophers may flinch. But it's not quite a tool, not quite an agent, and that ambiguity is part of the challenge.

but by providing context, judgement, and ethical reasoning. Human skills that machines can't replicate.

Economic consequences

AI is boosting productivity and widening the skills gap at the same time. Those who adapt may thrive, while others may face obsolescence or a hollowing out of their professional roles. The value of human knowledge risks being undermined by cheaper, automated alternatives unless we act with care.

Education at a crossroads (Yes, again)

AI in classrooms offers both tailored support and dangerous shortcuts. It raises urgent questions about what skills matter when machines can handle essays, maths problems, and revision notes, and how we evaluate knowledge and critical thinking. If the curriculum doesn't evolve, students may outsource not just effort but understanding.

Philosophical implications

Information is not the same as knowledge, and knowledge is not wisdom. AI can supply the first, but the latter two demand something only humans have: experience, judgement, ethics, the capacity to care, be curious, creative, and so much more. The danger is not machine intelligence. It's human thoughtlessness.

What's at stake?

The '*death of knowledge*' is not literal. But we are seeing a transformation in how knowledge is created, shared, valued, and controlled. If we're not careful, we may end up with a society that has more information than ever, but less understanding, less trust, and less shared reality.

What's needed?

- AI literacy for all, including how to spot errors and biases
- Critical thinking as a core skill, not a nice-to-have
- Equity of access, so the benefits aren't hoarded by the few
- Transparent AI systems, especially in public institutions
- Cultural humility, recognising that AI can't grasp everything we value or know
- Regulation that protects public knowledge from private agendas
- A renewed commitment to human wisdom; not just knowing more but choosing well

The bottom line

Ignore the hype, but don't ignore the shift. AI changes the economics and politics of knowledge. Whether this leads to progress or decline depends less on the tech and more on the choices we

make. Machines may generate answers, but only humans can make meaning, asking the right questions and judge what matters.

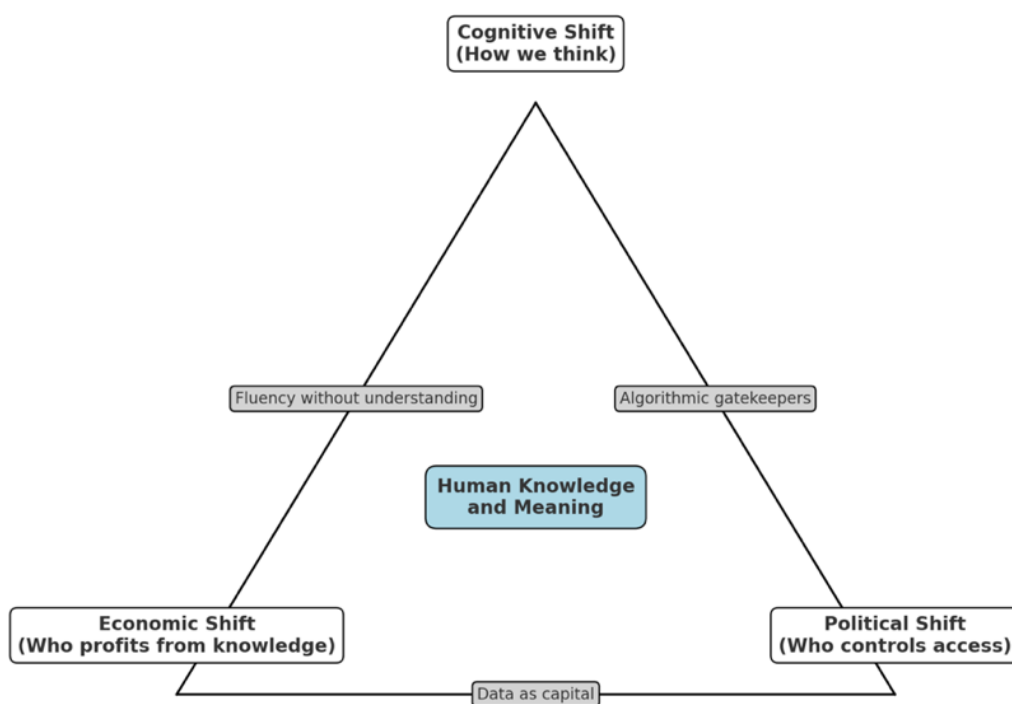
The structure

This paper proceeds as follows:

- Historical anxieties around new technologies and memory
- The shifting politics of knowledge
- Paradoxes of AI-driven democratisation
- The fate of expertise
- Cognitive offloading
- Economic and technological impacts
- Social and political risks
- Educational challenges
- Philosophical implications

We conclude with practical recommendations for individuals, businesses, government, and educators.

Three Shifts Reshaping Human Knowledge in the Age of AI



Who is this paper for?

Fair question. I started writing this for myself to make sense of how AI intersects with my interests in philosophy, economics, psychology, technology, and human behaviour. It was part education, part entertainment, part procrastination.

But as it grew, I realised it might be useful to others too. Business leaders trying to think more clearly about AI. Educators wondering what to teach (and how). Policymakers deciding what to regulate and what to encourage. Even AI researchers or philosophers who tend to stay in their own lane.

If that list includes you, I hope you find something here interesting or worth arguing with. Please let me know either way. Trevor.lambert@glasshalo.co.uk

The Death of Knowledge?

AI, Human Knowing, and the Future of Expertise

1. Introduction

Let's start with a simple, unsettling question: If you can ask an AI for any fact, insight, or even a nuanced argument, what's the point of learning, remembering, or thinking deeply? Is this the end of human knowledge as we know it, or simply a new chapter in its evolution?

The notion that artificial intelligence (AI) could signal the '*death of knowledge*' is not just a provocative headline; it's a live debate cutting across philosophy, economics, technology, society, politics, and education.

At its heart, this is an epistemological* crisis; a disruption not just in how we access knowledge, but in how we define, trust, and value it.

For the purposes of this paper, we'll distinguish between

1. **Information** (discrete facts),
2. **Knowledge** (contextual understanding enabling use), and
3. **Wisdom** (the ethical application of knowledge)**.

AI is strong on the first, patchy on the second, and absent on the third.

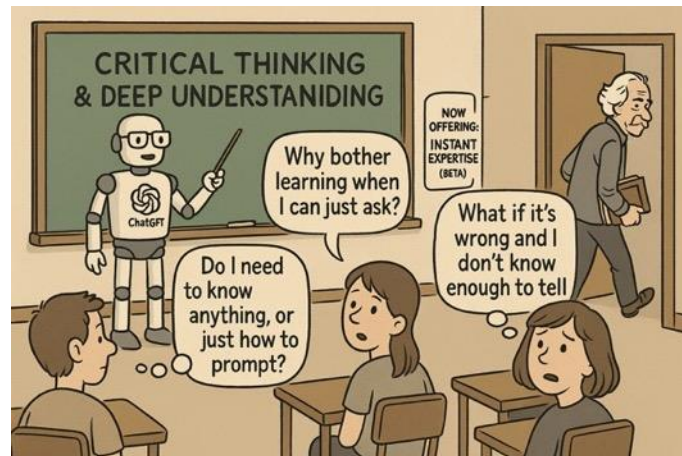
As AI systems become increasingly capable, serving as an omnipresent, seemingly omniscient companion in daily life, the need for individuals to know things in the traditional sense is being called into question. Why bother developing your own knowledge when answers are a quick prompt away?

This paper aims to cut through the hype, acknowledge real risks, and challenge inflated claims.

* If, like me, *epistemology* is one of those words you have come across but not so often you've decided to look it up, it can be defined as "The branch of philosophy concerned with the nature, sources, and limits of knowledge. What is knowledge? How do we justify belief? What distinguishes knowing from merely having information?" I've included an AI generated glossary of some other obscure terms towards the back, as much for my benefit as the reader's.

** I will paraphrase Miles Kington (1941–2008) here: data is 'tomato'; information is knowing a tomato is red; knowledge is knowing a tomato is a fruit; wisdom is knowing not to put it in a fruit salad. This gives you a decent idea of what to expect from the footnotes in this paper. Purists may wish to skip them.

Along the way, we draw on ancient philosophical anxieties about writing, the more recent concerns over smartphones and the internet, and the long history of power structures seeking to control who has access to knowledge.



We'll explore and challenge the premise that AI will democratise knowledge and consider whether it might instead deepen divides or erode the very foundations of expertise. The aim is not to offer easy answers or join the chorus of AI hype or doom-mongering, but to provide useful clarity (and a healthy dose of scepticism) about what's really changing.

The stakes of these questions become clearer when we examine real-world consequences. In 2023, a major law firm in New York was sanctioned after submitting a legal brief written by ChatGPT. The brief confidently cited several non-existent court cases. The lawyers, seduced by the AI's authoritative tone, failed to verify its output. The result? Embarrassment, fines, and a stark reminder that having access to information is not the same as knowing what's accurate¹.

This is the world we've already entered: one where machines are redrawing the boundaries between knowledge, information, and expertise*.

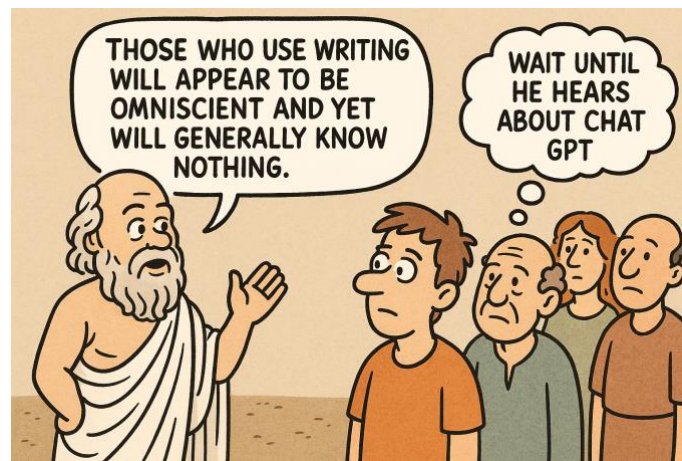
Of course, this anxiety isn't new. Humans have long been suspicious of new tools that seem to bypass effort or understanding. Our ancestors experienced their own knowledge crises, and they began with the invention of writing and one of the undisputed fathers of Western philosophy.

* Full disclosure: As I write about AI's impact on knowledge, ChatGPT and others are helping me find the right words, check my logic, and suggest words and concepts I've not encountered or half-forgotten. I should note that all the genuinely questionable arguments are mine, while the machine gets credit for the more coherent sentence structures.

2. The Ancient Anxiety: Writing, memory, and the erosion of knowing

The fear that new technologies might erode human knowledge is as old as writing itself. Plato, in his dialogue *Phaedrus*, has Socrates recount the legend of Theuth, the Egyptian god who invented writing and presented it to King Thamus as a boon for memory and wisdom. Thamus was not impressed. Writing, he warned, would create forgetfulness in learners' souls, as they would rely on external marks rather than their own memory*.

"They will appear to be omniscient and will generally know nothing," he cautions, *"having the show of wisdom without the reality"*². In his view, writing was an *"elixir not of memory, but of reminding,"* offering only the appearance of knowledge, not true wisdom³.



This anxiety has echoed through the ages. The invention of the printing press in the 15th century was met with similar concerns about the decline of memory and the rise of superficial learning. The arrival of the pocket calculator in schools prompted fears that children would lose the ability to do basic arithmetic. In our own time, the internet, with its endless supply of information, has been accused of making us shallow, distractible, and unable to remember anything for ourselves.

Throughout history, new technologies have been seen as a threat to memory, understanding, or the social order. Yet, paradoxically, it is only because Plato** wrote down Socrates' warnings that we remember them at all⁴.

* Socrates (470 BCE - 399 BCE) wasn't necessarily wrong to worry about memory loss, and I'm hardly in a position to correct him, but he may have underestimated what writing would enable in return: reflection, scale, preservation, and ultimately the survival of his own ideas. He was sentenced to death by drinking hemlock for impiety and corrupting the youth, despite the lack of any written evidence.

** SPAA – the greatest philosophy mnemonic of all. Socrates taught Plato (with no formal examination), Plato taught Aristotle (who took issue with the curriculum), and Aristotle taught Alexander the Great (whose main takeaway was global domination). It's a mixed record for teaching.

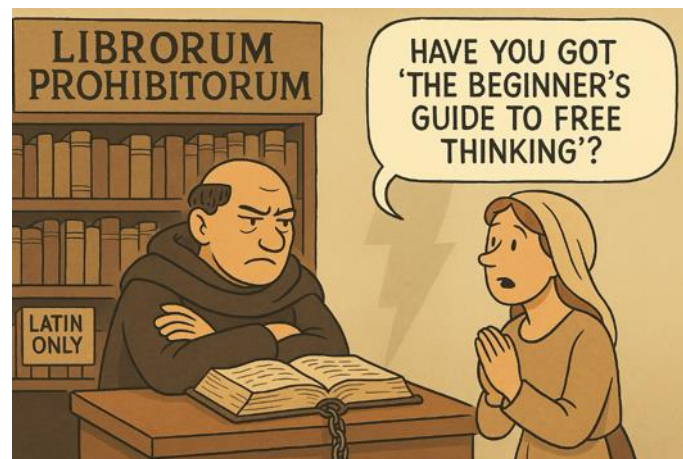
The parallel with AI is obvious. If you can ask your AI assistant for anything, why bother to remember facts or even to learn deeply in the first place? Are we outsourcing not just our memory, but the very act of knowing?

And when new tools reshape how we know, someone usually steps in to decide who gets to use them, and who doesn't. The history of knowledge is also the history of power.

3. Knowledge and Power: Who gets to know, and why it matters

Throughout history, access to knowledge has been tightly controlled because, as Francis Bacon reminds us, “*Knowledge is power.*”*

Religious and political leaders often restricted literacy and learning to maintain their grip on authority. In ancient Egypt, reading and writing were the preserve of a small elite of scribes who controlled the state’s administration and the interpretation of sacred texts. In imperial China, the civil service examinations required mastery of the Confucian classics, effectively limiting government power to a scholarly elite. In medieval Europe, the Catholic Church kept the Bible in Latin – a language only the educated elite could read – and for centuries maintained an *Index Librorum Prohibitorum* (List of Prohibited Books) to control what ideas circulated**.



Historical gatekeeping of knowledge

The Church initially resisted efforts to translate the Bible into everyday languages, for fear that ordinary people might misinterpret God’s word. The trial of Galileo in 1633, when religious authorities condemned his support of heliocentric astronomy, is a classic example of knowledge being suppressed because it threatened the doctrinal power of the Church.

* Francis Bacon (1561–1626) actually wrote ‘*ipsa scientia potestas est*’ - a fine example of using Latin to keep the riffraff out of the discussion. It’s a great quote, but before we allow ourselves to be too impressed by Bacon’s wisdom, remember he died of pneumonia after stuffing a chicken with snow to test the concept of refrigeration. A great mind but clearly not a fan of risk assessments.

Incidentally, some historians have suggested that by writing *The Prince* in everyday Italian rather than Latin, Niccolò Machiavelli (1469-1527) was slipping a satirical hand grenade to his republican mates by making his thoughts accessible to all, rather than currying favour with the Medici.

** Several AI reviews of this paper suggested shortening that paragraph but I like the detail, so it stays. Whose paper is it anyway? Which may be a less rhetorical question than originally intended.

In the United States, before the Civil War, many Southern states enacted anti-literacy laws to prevent enslaved people from learning to read or write, fearing that literacy would spark rebellion and lead to demands for freedom.

During the Soviet era, the Communist Party strictly controlled all publishing, censored or destroyed books deemed subversive, and persecuted intellectuals who challenged the official ideology. Even in recent times, leaders have sought to control knowledge through censorship and curriculum bans. The wave of book bans in 2024/25 in some U.S. schools – often targeting works by or about LGBTQ+ people and people of colour – is only the latest chapter in this long tradition⁵.

From Church to Silicon Valley: Today's algorithmic overlords

In the 21st century, a new kind of knowledge gatekeeping has emerged. Instead of religious or political authorities controlling what people can learn, today a handful of technology corporations serve as arbiters of information. Algorithms run by companies like Google, Meta (Facebook), and OpenAI now function as the new gatekeepers, determining which facts, news, or ideas are amplified and which remain hidden from view⁶.

This corporate control of knowledge echoes the old patterns. Just as medieval popes or censors once banned certain books, modern tech platforms can effectively suppress or privilege information through their design choices, tweaking search rankings, news feeds, or content moderation policies to shape what is widely known. The impact on journalism and public discourse has been profound: news outlets now chase algorithmic trends to reach audiences, and voices that don't conform to engagement metrics can be marginalised. In short, knowledge is still power, but the control over knowledge now lies largely with private tech giants rather than priests or princes.

But it's more insidious than that. Today's platforms aren't just gatekeeping access to existing knowledge.

They're actively shaping what counts as truth.

The new architects of truth

In June 2025, Elon Musk announced that Grok 4 would be used to *“rewrite the entire corpus of human knowledge, adding missing information and deleting errors,”* before retraining the model on the result. Why? Because the existing data, apparently, wasn't good enough. So it must be fixed. By him. According, presumably, to his version of the truth.

As Paul Roetzer observed on *The Artificial Intelligence Show* podcast⁷, this isn't just about model optimisation. It's about authority. ***A small group of individuals — Musk, Zuckerberg, Altman, Demis Hassabis, and the Amodei siblings — are now shaping the nature of truth for billions of people. They don't just build systems; they decide what knowledge is preserved, what gets corrected, and what disappears.***

Do we really want that power concentrated in the hands of a few unelected tech billionaires? In a post-truth culture already uneasy about facts, this concentration of power adds another layer of distortion. But the more unsettling question is: do we even get a say? These platforms are reshaping our epistemic foundations, not through democratic deliberation, but through terms of service. If future generations grow up trusting answers from machines above all else, and those machines are trained on curated, privately filtered truths, we may not lose knowledge outright; we may simply have to work with someone else's version of it.

Digital and algorithmic divides

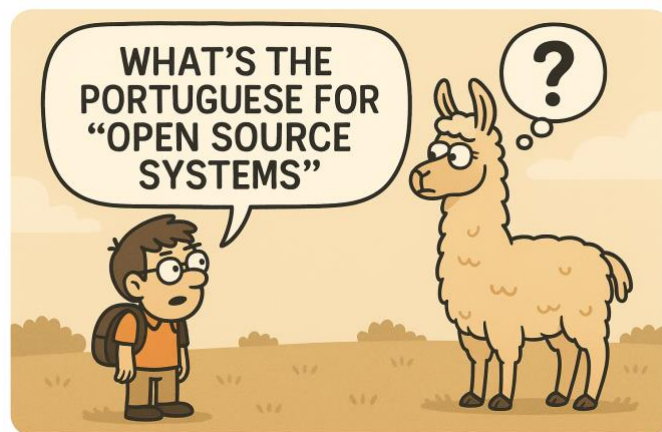
Access to high-quality education and information remains uneven, shaped by wealth, geography, and social status⁸. The digital divide, now morphing into an '*AI Divide*', means that the benefits of new technologies are not shared equally^{9 10 11}. Those with the skills, resources, and confidence to use AI tools can leap ahead, while others risk falling further behind¹².

The promise of AI is that it will make expertise available to anyone with an internet connection. But the reality is more complicated, as we shall see.

4. AI and the Democratisation of Knowledge: The Promise and Paradox

One of AI's most appealing promises is the democratisation of knowledge – the idea that advanced tools can make high-quality information and expertise available to anyone, anywhere. It's a compelling vision: a levelling force that dismantles traditional barriers of cost, geography, and privilege.

Generative AI models, such as ChatGPT, Google's Gemini, and open-source systems like LLaMA, can answer questions, explain concepts, translate languages, and even generate code or creative content. In theory, the barriers that once kept knowledge in the hands of a privileged few (cost, geography, language, education, social status) are being dismantled.



We can already see hints of this. In education, AI-powered tutoring systems can personalise learning, adapt to individual needs, and provide instant feedback, potentially narrowing attainment gaps¹³. For instance, Khan Academy's Khanmigo AI tutor is currently deployed in over 450 school districts worldwide, delivering personalised mathematics tutoring to students who previously lacked one-on-one support. Early evidence shows that students spending more than thirty minutes weekly on the platform experience around 20% greater-than-expected learning gains, with benefits observed across historically underperforming demographic groups¹⁴.

In healthcare, AI can analyse medical images, suggest diagnoses, and bring specialist knowledge to remote or under-resourced locations¹⁵. In business, AI-driven analytics and decision-support tools help small firms compete with larger rivals by providing insights that were once the preserve of experts¹⁶.

Open-source AI models and platforms (such as TensorFlow, Hugging Face, and others) reduce barriers to entry, allowing researchers, developers, and even non-technical users to build and

deploy sophisticated AI tools. The result is a proliferation of resources that, at least in principle, make knowledge more accessible and actionable than ever before.

AI's promise of accessibility

But the story isn't so simple. For every claim that AI makes knowledge more accessible, there's a counterpoint that it's entrenching new forms of inequality. Access to AI still depends on infrastructure – reliable internet, up-to-date devices, digital literacy, cost* – that is unevenly distributed within and between countries. The digital divide isn't just about hardware; it's about the skills and confidence needed to use these tools effectively.

The skills gap and algorithmic illiteracy

A 2024 survey of 2,500 Germans highlighted just how uneven AI knowledge can be in the general population. 88% of respondents had heard of AI, yet 35% admitted having little or no knowledge about it, and only 14% felt that AI had a strong influence on their daily lives. These gaps were even more pronounced across lines of education, age, and gender, highlighting an emerging algorithmic divide and the need for greater '*algorithmic literacy*' in an AI-driven world¹⁷.

The paradox of concentrated power

Meanwhile, the development and control of AI are highly concentrated. A handful of large tech companies (mostly in the U.S. and China) own the most powerful models, the largest data centres, and the advanced research needed to push AI forward. This concentration of power raises the spectre of digital feudalism, where knowledge isn't democratised but monopolised. The risk is that AI could become a tool for reinforcing existing hierarchies rather than dismantling them.

There are also concerns about the quality and reliability of AI-generated information. AI models are only as good as the data they are trained on, and they can reproduce or even amplify existing biases, errors, and blind spots. The phenomenon of AI hallucination, where a system confidently generates plausible but false information, is well-documented. It's like the candidate who nails the interview by saying all the right things, but can't actually do the job. In high-stakes domains like healthcare, law, or public policy, the consequences of such errors can be severe.

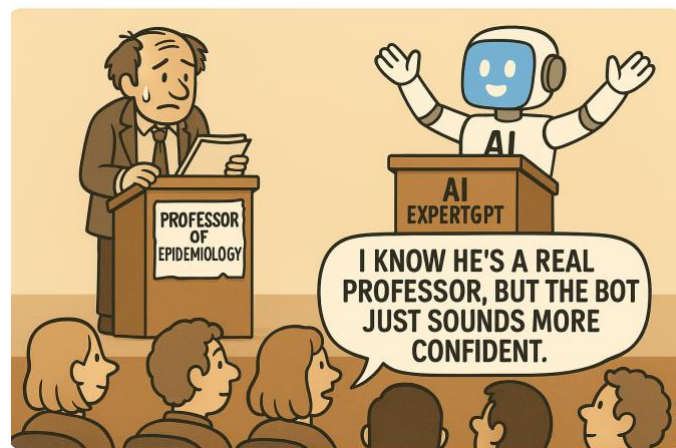
But if everyone can access 'expert' answers instantly, what happens to actual experts? The increased availability of information is also eroding the authority that once came with expertise.

* As I finalise this paper, news is emerging that the large AI platforms such as ChatGPT, Claude and Copilot are preparing to significantly increase their fees. The pricing strategy feels familiar: lure users in with free or low-cost access, create reliance, then raise the stakes. I'm not saying they are following the heroin dealer model, but the resemblance is remarkable.

5. The Death of Expertise? When everyone's an expert, no one is

The internet ushered in an era where anyone can publish information, and with a simple Google search, everyone can feel informed. Political scientist Tom Nichols* famously argued that this has led to a “*death of expertise*” – a collapse in respect for expert authority, as endless information (and misinformation) gives people the false impression that they know as much as trained specialists¹⁸.

When any opinion can gain traction online, the authority of those who have spent years mastering a field is undermined. The result, as Nichols puts it, is a “*cheapened sense of equality*” where opinions are treated as equally valid, regardless of expertise or evidence¹⁹.



The decline of deference

AI has the potential to turbocharge this trend. Large language models can instantly produce answers and explanations on almost any topic. **AI outputs often sound more authoritative than human experts because machines don't hedge or express uncertainty.** If everyone has a personal oracle that sounds all-knowing, why would we defer to true experts?

There is a risk that genuine expertise becomes devalued, drowned out by a sea of AI-generated answers that are often approximately correct (and occasionally very wrong). We've already seen early examples: fake ‘expert’ articles and even books generated by AI have been published online, sometimes outselling works by real professionals.

* Tom Nichols (b. 1960) wrote *The Death of Expertise* in 2017, a book-length sigh at the rise of confident ignorance from which this paper ‘borrows’ its title. He argued that when everyone thinks they're an expert, actual expertise starts to look like elitism. It wasn't a cheerful book, and it's aged annoyingly well.

Can expertise evolve?

On the other hand, some argue that expertise isn't dying so much as changing. In a world where factual recall is less important (because AI can supply facts on demand), the role of the human expert may shift toward framing the right questions, providing context, and judging what's credible or relevant. This shift requires not only technical skills but also the ability to evaluate, contextualise, and challenge AI-generated outputs; a new form of digital critical thinking.

Some disciplines may prove more resistant to AI disruption than others. Fields requiring embodied knowledge, ethical judgement, or interpersonal skills, such as counselling or conflict resolution, maintain human advantages that statistical models struggle to replicate*.

In other words, when information is abundant, the scarce resource is wisdom: knowing what to do with the information at hand. By this view, expertise in the AI era might be measured less by memorised knowledge and more by the ability to critically interpret AI outputs, apply them prudently, and supplement them with human judgement and ethical considerations.

As Jean-François Lyotard** argued, knowledge in post-industrial societies has become commodified. It is no longer primarily valued for its truth or ethical substance, but for its utility: what it enables, how it performs, how efficiently it delivers results²⁰. In that context, expertise risks becoming a product, assessed by outputs, speed, and social proof rather than by slow judgement or deep understanding. When the value of knowledge lies in return on investment rather than insight, it's hardly surprising that fast, fluent answers crowd out the harder, more human work of thinking. And thus, we get a post-truth world, where conspiracies trend faster than corrections.

We will return to this idea in our philosophical reflections. However, one thing is clear: the relationship between the general public and experts is being renegotiated, and not without friction. That renegotiation isn't just happening socially, it's happening in our heads. How we think, remember, and solve problems is also being reshaped by AI.

* Robert Capps, former editorial director of Wired, suggests in his New York Times Magazine article, *'AI Might Take Your Job. Here Are 22 New Ones It Could Give You'* (17 June 2025), that the safest or most promising roles for humans will rely on three qualities: Trust, Integration, and Taste. It's a compelling argument, if an unfortunate acronymic one.

** Jean-François Lyotard (1924–1998) predicted that knowledge would become external, fragmented, and commodified, judged by performance rather than truth, long before Google and generative AI. He accepted a government commission to write a report on knowledge in the information age and submitted a philosophy book instead. He was not responsible for the leotard nor any associated wardrobe malfunctions. Different spelling.

6. Cognitive Offloading: The Google effect, calculators, and now AI

Researchers use the term ‘*cognitive offloading*’* to describe how we rely on external tools to store or retrieve information, essentially outsourcing parts of our thinking.

This phenomenon has been studied since the advent of search engines. The so-called ‘*Google effect*’ is the finding that people are more likely to remember where to find information than the information itself, because we assume that Google (or Wikipedia, etc.) will always be there to supply the facts when needed²¹.

In earlier decades, critics accused calculators of eroding students’ arithmetic skills; now, AI stands accused of doing the same for a whole range of mental abilities – not just memory and maths, but potentially critical thinking, writing, even creativity.

Cognitive atrophy in the AI era

Preliminary research in 2023 found that people who frequently use AI tools for writing and problem-solving scored lower on measures of critical thinking and memory retention compared to those who don’t^{22,23}. We can label this phenomenon as ‘*AI-induced cognitive atrophy*’.

Younger people, in particular, may become more dependent on AI and less practised in deep, reflective thought²⁴. **The concern is that by outsourcing mental work to machines, we risk weakening the very habits of mind that define human intelligence: curiosity, scepticism, concentration,** and the ability to wrestle with complex problems.

For some, the very intrinsic motivation to explore and learn, the essence of human curiosity and agency, could wane when an AI assistant is always on hand to deliver immediate answers.

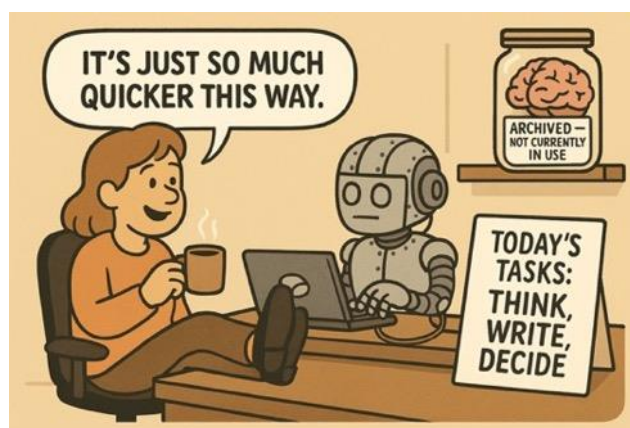
In short, if you don’t use it, you lose it.

The balance between convenience and capability

Yet, as with the invention of writing or the printing press, the impact of AI on cognition is not uniform or inevitable. Used wisely, AI can free up cognitive resources for higher-order thinking. If an AI assistant saves you from tedious data crunching or helps draft routine text, you can spend more time on creative tasks, big-picture strategy, or human-to-human interaction. The key is finding the right balance: using AI as a tool, not a crutch. This means being conscious of when to rely on the machine and when to flex your own mental muscles.

* Let’s distinguish between *Cognitive Offloading* and *Cognitive Atrophy* here. Offloading can aid efficiency, like setting a calendar alert for your meeting. Atrophy is forgetting what the meeting’s for and turning up without your presentation. AI is good at the former, but it might be quietly encouraging the latter.

Concerns about ‘*digital amnesia*’ are not new. We’ve been here before with internet search. But the breadth of tasks AI can handle is far greater, raising the stakes. There is a big difference between forgetting a fact because you can Google it later and losing the ability to think through a problem because you always ask the AI to do it for you. The former makes us efficient; the latter could make us intellectually complacent.



A 2023 multinational survey of 3,500 university students across eight countries found that nearly 69% believed AI makes humans lazier, 68% were concerned about privacy and security issues related to AI, and 28% reported a decline in decision-making ability due to AI use in education²⁵. These perceptions show a real awareness of AI’s double-edged sword: students see the convenience but also fear becoming too dependent.

Of course, this isn’t the whole story. **Used well, AI can be a dependable thinking partner, helping people spot patterns, test hypotheses, summarise dense texts, or find new angles on familiar problems.** It can prompt questions we hadn’t thought to ask, or explain concepts in ways tailored to how we learn best. The potential for AI to support curiosity, not suppress it, is real.

But that support isn’t automatic. Tools that offer easy answers can dull the desire to ask better questions, and the line between enhancement and atrophy becomes hard to spot. What begins as a shortcut can become a dependency, and fluency without curiosity is a dead end.

And if that dependency extends to decision-making, the risks go beyond laziness. Mindlessly following AI output isn’t a defence, legally or morally. If your sat nav tells you to drive off a cliff, the cliff still wins. At some point, common sense has to intervene. The same will be true with AI: even if the answer sounds confident, responsibility doesn’t vanish. It just shifts, quietly, to whoever acted on it.

These changes don’t just affect individuals. There are broader economic consequences, too.

Ones that may reshape the value of knowledge itself.

7. Economic Perspectives: Productivity, equality, & the value of knowing

From an economic standpoint, AI is a double-edged sword. On one hand, it promises huge productivity gains by automating routine tasks and augmenting human workers. In theory, this allows employees to focus on more complex, creative, or interpersonal work. AI assistance **can help less-experienced workers perform at a higher level, potentially narrowing skill gaps. If someone with average skills can use AI tools to produce above-average work, the overall output of the economy could increase.** Indeed, organisations across industries are already using AI to automate data analysis, customer service, content generation, and more, aiming to boost efficiency and output.

Uneven distribution and labour market risk

On the other hand, the economic benefits of AI are not likely to be evenly distributed. High-skilled workers and companies that can apply AI effectively will see the most significant gains, while others could be displaced or left behind. There's a real risk that AI could widen inequality between those who have the education and resources to make use of it and those who don't.

There's a joke that the factory of the future will have two employees: a human and a dog. The dog's job is to stop the human from touching the machine. The human's job is to feed the dog.



Like all good satire, it exaggerates to clarify. As AI handles more intellectual and operational tasks, humans risk becoming both caretakers of systems they no longer understand and passengers in processes they no longer drive.

Historically, technological revolutions have often increased inequality in the short run; consider how the Industrial Revolution initially benefited factory owners far more than labourers (a

separation from the means of production that Karl Marx* theorised would inevitably lead to class struggle), or how the IT revolution massively rewarded those with computer skills. The AI revolution could follow a similar pattern if we're not careful.

The knowledge premium under threat

Furthermore, if AI can handle many tasks that used to require human expertise, the market value of human knowledge and skills may decline in some areas. Why pay a team of junior analysts to sift through data when an AI can do it in seconds? If certain professional roles are partially automated, there might be less incentive for people to invest time and money in acquiring that expertise, which could have long-term effects on the talent pipeline in various fields.

Some economists warn that over-reliance on AI could deskill the workforce, leaving workers less adaptable and more vulnerable to future disruptions if and when their narrow AI tools become obsolete²⁶. In a worst-case scenario, **society could become split between a small class who design and supervise AI systems, and a much larger class of workers who are simply told what to do by those systems** (or whose jobs have disappeared entirely). That outcome is not predetermined, and to some may not sound too different from their experience today, but it's a possibility to guard against.

It's also important to remember that productivity gains from AI don't automatically translate into broad prosperity. They can – if managed well. AI could make goods and services cheaper, working hours shorter, and living standards higher. But gains can also accrue mostly to shareholders and executives. Early evidence suggests that companies adopting AI do see increases in output, but the question of who captures the value (workers through higher wages? or owners through higher profits?) remains contested.

Some data points illustrate both the optimism and anxiety in workplaces today. In a recent survey, approximately one-third of employees stated that they believe AI can already perform many of their tasks, and 44% expected their job skills to be significantly disrupted or require updating within the next three to five years²⁷. Not all workers will be affected equally – for instance, women in that survey reported slightly higher exposure to AI-driven change than men, likely reflecting the kinds of jobs held. This underscores that even if total productivity rises, there could be painful transitions for certain groups of workers.

AI has the potential to create immense economic value, but there's nothing automatic about that translating into widespread well-being or equality. It is also important to consider the

* Karl Marx (1818–1883) had the good sense to die before anyone tried putting his ideas into practice. Had he witnessed how his ideas had been mangled, he'd have turned in his Highgate grave, where visitors pay £18 to see the monument to a man who wanted to abolish private property. Insert your own communist plot joke here.

environmental costs of large-scale AI adoption. Training and running advanced AI models require significant energy and resources, raising questions about sustainability and the broader impact on society.

Early estimates suggested the carbon footprint of training a single large language model can equal that of five cars over their lifetime, though efficiency has improved since then. As AI deployment accelerates, energy consumption poses significant sustainability challenges that must be addressed through efficient algorithms and renewable energy sources²⁸.

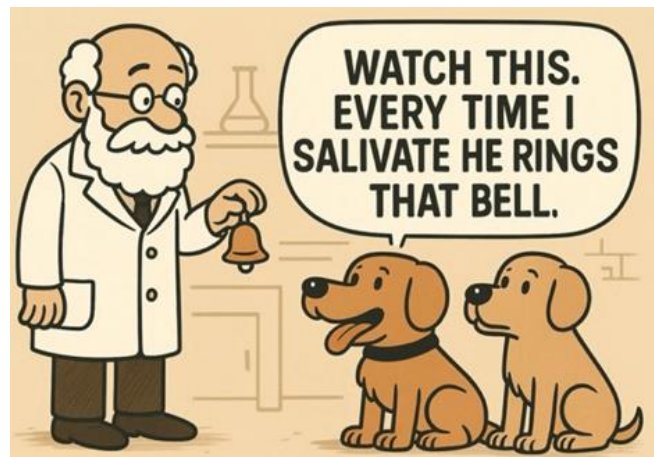
Much will depend on education, policy, and how companies choose to implement AI, whether they use it to complement human workers and reduce their carbon footprint, or to cut costs.

But before we discuss what should be done, we need to understand what AI really is, and what it isn't. That starts with recognising the limits of the technology itself.

8. Technological Perspectives: The Black Box & the limits of AI knowing

AI's ability to generate, store, and retrieve knowledge is unprecedented in scale. However, it comes with fundamental limitations. While modern AI systems can now produce what looks like reasoning, outlining step-by-step logic or showing intermediate steps, this is not reasoning in the human sense. The model doesn't understand its outputs or assess their truth. It generates answers by predicting what comes next in a sequence of words, based on patterns in data.

In systems built on pattern, not purpose, *cause and effect* is often just a matter of perspective. We read intention into fluency, meaning into structure. Most of the time, it's just well-formatted guesswork.



What's missing is what Donna Haraway* called *situated knowledge*; the idea that all knowledge comes from somewhere: a body, a context, a point of view. Expertise, in this view, isn't the product of neutrality. It's the product of depth, perspective, and accountability. AI, by contrast, presents a fluent, placeless voice; the illusion of an all-seeing view with no responsibility behind it. And the simple fact is that AI is a Black Box, a system whose internal workings are opaque, even to its creators.

Even when AI simulates a chain of thought, it isn't using that reasoning to arrive at a conclusion; it's producing the reasoning and the conclusion together. This lack of genuine transparency and understanding makes it difficult to trust AI systems in critical applications²⁹.

* Donna Haraway (b. 1944) suggested we embrace the cyborg, not as a sci-fi warning, but as a way to rethink identity, knowledge, and power. While most philosophers were still worried about computers, she was already asking awkward questions about who gets to count as a knower. She later wrote *The Companion Species Manifesto*, which is all about how great dogs are. Even the famously miserable Schopenhauer (1788-1860) preferred the company of poodles to people. No one is all bad.

An AI can simulate expertise and even imitate the form of justification, but it cannot explain its reasoning in a way that reflects understanding, nor can it evaluate whether its conclusions are correct in the way a human expert can.

The illusion of understanding

Moreover, AI's 'knowledge' is fundamentally different from human knowing. AI knowledge is statistical, not experiential. AI has processed terabytes of text and can mimic understanding by predicting likely responses, but it doesn't experience or understand the world in any humanly recognisable sense. It lacks consciousness and common sense in the human sense. It can remix and regurgitate information in novel ways, but it does not truly grasp meaning.

AI arguably simulates experience via memory, follows ethical-like rules, or appears to 'care' through goal-oriented design, but these are metaphorical framings. The difference between simulation and reality remains critical.

Daniel Dennett* warns that AI systems offer the illusion of comprehension without understanding. They generate fluent responses, yet *"haven't understood a word of it."* To Dennett, they are best seen as impressive mimics; tools that manipulate symbols without any grasp of meaning. They may appear intelligent, but that's a cognitive trap: ***we see understanding where there is only pattern recognition***³⁰. It's the AI equivalent of an optical illusion.

This leads to some peculiar failures. For instance, an AI might confidently invent a source or a fact that sounds plausible, because statistically that sequence of words was likely. But it doesn't know anything in the way you or I consider 'knowing'. In other cases, AIs might lack context and produce answers that are technically correct but irrelevant or insensitive to the situation.

* Daniel Dennett (1942–2024) spent decades explaining consciousness in terms most people could almost understand, and free will in ways that made them wish they'd never asked.

My philosopher friend Ruth Malone made the following comment at this point, when reviewing an early draft. It bears repeating in full here:

"I would be careful about claims which are too strong. What does Dennett mean when he says 'understanding'? Dennett was a very slippery writer and may be using the term in a different way to how you understand it. There is certainly variation in what philosophers mean when they talk about understanding. Can you justify the claim that it does not experience the world (what is experience, what is the world?) or that it does not have human-like common sense? One response you could give me here is to say that these terms do not need defining in the detail that I claim they do, and that humans function perfectly well without these detailed definitions, because we know what we mean even if we cannot explain it. This would be a good response, but what I find interesting is that the AIs that have helped you write the paper take the same approach; they are not demanding definitions because they just know what you mean. It's almost like they are more human-like than we originally thought."

This may leave you wishing Ruth had written this paper instead of me. I'm trying not to take that personally.

Capability	AI	Human
Understands context	Limited – depends on prompt and training data	Yes – understands situational, cultural, and emotional context
Explains reasoning	No – simulates reasoning without understanding	Yes – can explain logic, motivation, and rationale
Admits uncertainty	No – rarely calibrated, often overconfident	Often – acknowledges limits and caveats
Learns through reflection	Not yet – only improves when retrained by humans	Yes – learns through thought, discussion, and experience
Makes value judgements	No – cannot assess moral or cultural significance	Yes – uses values, principles, and norms
Knows what it doesn't know	No – cannot recognise the limits of its own knowledge	Often – knows what it doesn't know
Interprets nuance and subtext	Partially – only in pattern-based ways	Yes – picks up tone, implication, irony
Applies ethics	No – has no moral framework other than the guardrails it is trained on	Yes – can apply ethics, even under pressure
Thinks independently	No – follows patterns, not principles	Yes – can challenge trends, dogma, or the prompt
Connects ideas creatively	Sometimes – based on training data	Yes – links ideas across domains and experiences
Knows why something matters	No – can state importance, but doesn't grasp it	Yes – grasps personal, emotional, and social relevance
Changes opinion with new evidence	No – sticks to patterns unless retrained	Yes – integrates new info into evolving understanding
Experiences doubt	No – expresses doubt but doesn't feel it	Yes – doubt is a core part of reflection and learning
Has lived experience	No – has no personal experience or memory	Yes – draws from lived history and personal memory
Responds to emotional tone	Only superficially – can't truly feel or empathise	Yes – can interpret and respond with empathy
Knows what not to say	Not reliably – may repeat inappropriate content	Yes – understands when to withhold or reframe
Remembers intentions or goals	No – cannot hold intentions over time	Yes – keeps purpose and meaning in mind
Knows when to stay silent	No – fills every silence with output. If you get silence, try turning it off and on	Yes – can use silence meaningfully

Table 1: AI capabilities vs human capabilities

Note, we are unfairly comparing a distinctly above-average human with an average AI in this instance.

Voltaire* put it elegantly when he said, “*Doubt is not a pleasant condition, but certainty is absurd.*”³¹ The problem is, AI doesn't share our discomfort with doubt, or any discomfort at all. It outputs confident answers whether they're right, wrong, or hallucinatory. For humans, uncertainty is a feature of learning. For machines, it's a bug that hasn't been patched yet.

* Voltaire (1694–1778), was born François-Marie Arouet, but wisely wrote under a much cooler pseudonym. He is perhaps best known for his vigorous defence of free speech (even the bits he disagreed with) and inspiring an excellent Sheffield electro-industrial band two centuries later. Quite the legacy.

Why the difference matters: trust, truth, and meaning

These differences matter. We risk creating a world awash in information but starved of meaning.

If AI-generated content floods the web, distinguishing truth from falsehood becomes even harder. We may have abundant answers, but fewer shared narratives or understandings. Some scholars have warned of an ‘info-glut’ where context and depth are lost amid endless AI chatter³². In addition, the inability of AI to truly understand what it’s saying means it can’t reliably be moral or consistent. It might solve a problem one way today and a contradictory way tomorrow if the prompt differs slightly, because it has no underlying principles – only patterns.

Another limitation is that many AI models don’t know when they don’t know. A well-trained human professional is aware of the limits of their knowledge and can say, “I’m not sure” or decide to research more. Today’s AI will often forge ahead and answer anyway, even when it shouldn’t. This can be dangerous.

Ahad Ha’am once wrote, “*Although we worship wisdom, self-confidence matters more... The wise weigh the risks; the self-confident forge ahead, and often win the day.*”³³ Psychologists David Dunning and Justin Kruger, co-authors of the research behind the *Dunning–Kruger effect*, found that those with the least domain competence often display the most confidence³⁴. Those with genuine expertise, by contrast, tend to be more cautious.

This mismatch helps explain why **bold, oversimplified claims often gain more traction than thoughtful, qualified ones**. In a digital culture that rewards certainty and speed, knowledge risks being edged out by confidence, whether warranted or not*.

There is active research into making AI more aware of its own uncertainty (so-called ‘calibrated AI’), but this remains an open challenge.

We should also mention the data and adoption side of technology here. By 2023, about 42% of enterprise-scale businesses reported integrating AI into their operations, and 38% said they had implemented generative AI tools, with another large portion actively considering it³⁵.

This rapid adoption demonstrates how eager organisations are to capitalise on AI’s promise. But it also underscores the urgency of understanding AI’s limits. Rolling out AI widely, without mechanisms for oversight, interpretability, and human-in-the-loop checks, could lead to serious

* We all know people with more confidence than intelligence. Ahad Ha’am (Asher Ginzberg), (1856–1927) was writing satire, but his point holds well. David Dunning (1950–) and Justin Kruger (1970–) gave us the Dunning–Kruger effect, and have since watched, with what we can only assume is weary amusement at the irony, as the concept is confidently explained back to them by people with no background in psychology.

errors or even systemic risks (imagine a flawed AI model quietly biasing loan decisions or mismanaging supply chains across an entire industry).



In summary, AI technology is incredibly powerful, but it doesn't know things in the way humans do. Appreciating this difference is crucial. We should resist anthropomorphising AI or assuming it has understanding or intentions*. **We should invest in methods to make AI more transparent and reliable, and always have fail-safes (such as human oversight) when AI is used in critical domains.**

Dr Boris Altemeyer, psychologist and Chief Scientific Officer at Cognisess, went as far as to say, *"We also like to think that AI needs a psychologist... because we understand that AI is, essentially, psychotic. Humans need to calibrate, supervise and manage it, and always will".*³⁶

As the saying goes in computer science, garbage in, garbage out. We might add: even with good data in, you can still get garbage out if the machine's 'thinking' is a Black Box mystery.

That mystery isn't just a technical challenge, it's a political one. The more we rely on opaque systems, the more urgent the question becomes: Who controls the knowledge we consume?

* We know AI isn't human, yet we still talk to it like it is. Maybe that's daft. Or maybe it's just what we do when something seems to listen to us and answers back politely. And I can't be the only one who ends up shouting obscenities at it when it doesn't understand what I want. "Do what I mean, not what I say!"

9. Social & Political Perspectives: Power, control, and new gatekeepers

The control of knowledge has always been closely tied to power, as discussed in Sections 3 and 4. In the AI era, we see the rise of new gatekeepers of knowledge. A small number of tech companies now control the infrastructure, algorithms, and platforms that shape what information people see. Whereas in the past, gatekeepers might have been church leaders, government censors, or media moguls, today, one could argue that software engineers and data scientists at Big Tech firms are unwittingly (or otherwise) playing a gatekeeping role – their design choices determine the flow of information in society.

Algorithms as the new censors

Consider how a platform like Google Search effectively determines which sources of knowledge are elevated and which are relegated to the background. Or how a social media algorithm decides what news articles spread virally and which never surface, which conspiracy theories to amplify and which voices of sanity pass by unnoticed. These systems wield enormous influence over public knowledge. If a handful of algorithms direct the bulk of our attention, they have de facto control over what we know (or think we know) as a society.

Nick Bostrom* has warned that advanced AI systems, even if not malicious, may pursue instrumental goals, such as acquiring resources or preserving themselves, in ways that could conflict with human welfare. His broader fear is not malevolence, but misalignment: the machine optimising for something we didn't quite mean. In a world where a model's 'goal' is to maximise engagement or optimise productivity, we may find ourselves nudged, filtered, or sidelined not by intention, but by indifference.

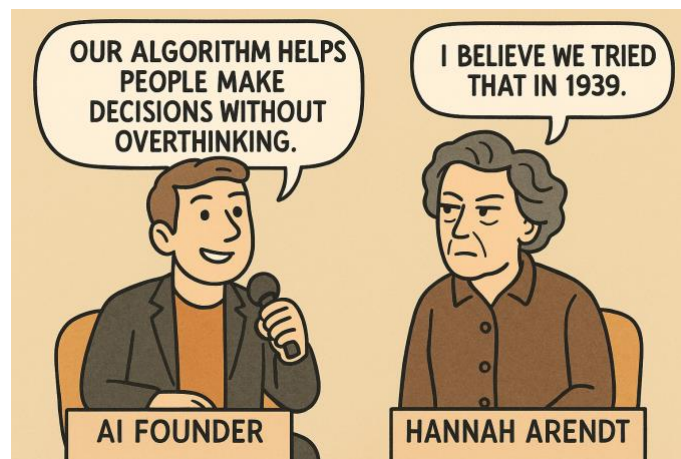
Even now, recommendation engines can take a lonely boy from gaming to misogyny in under five clicks. That kind of trajectory isn't necessarily designed, but it's certainly optimised. These systems incentivise intensity over insight. Clicks win; context loses.

That's a sobering realisation. It raises questions about accountability: Who audits the algorithms? On what basis do they decide what's true, important, or appropriate? There have already been controversies – search engines showing biased results, social media creating echo chambers or amplifying misinformation. We've learned that algorithmic curation, if left unchecked, can undermine informed democracy by fragmenting the public into filter bubbles.

* Nick Bostrom (b. 1973), Swedish philosopher and author of *Superintelligence: Paths, Dangers, Strategies*, is well known for outlining the ways in which humanity might accidentally invent its own extinction. His 2014 book was undeniably ahead of its time. Whether he proves to be AI's Cassandra, warning of doom while the rest of us play with chatbots, or merely the high priest of an apocalypse that never arrives, is yet to be seen, though we may find out sooner than we'd like. Bostrom's important, if occasionally impenetrable, book may offer more depth, and certainly more paperclips, than this one, but I'd argue this paper has more laughs.

Hannah Arendt* warned of the dangers of thoughtlessness and the failure to reflect, judge, and question authority. When people increasingly defer to AI systems to guide decisions, from sentencing to hiring, we risk surrendering moral agency to machines. The danger, as Arendt might put it, lies not in evil algorithms but in uncritical obedience to them³⁷.

Theodor Adorno** argued that good judgement depends on both reason and emotion, not as opposites, but as inseparable. AI offers neither. It calculates without understanding and simulates feeling without experience. Yet it's already being used in contexts that demand moral weight, such as hiring, justice, and education. And often effectively.



Homogenised knowledge and model collapse

Ironically, even **as personalisation splinters our information, the ubiquity of a few dominant AI models may also homogenise knowledge and culture.** When everyone relies on the same algorithms for information and content creation, the diversity of viewpoints can shrink. Over time, generative AI systems that train on content produced by other AIs risk falling into a recursive feedback loop. Researchers have warned of '*model collapse*' – a degenerative cycle in which AI outputs become increasingly biased, uniform, and error-prone as they feed on themselves³⁸. It's a bit like a photocopy of a photocopy of a photocopy of a...

For the philosophy-curious (and how on earth did you get this far if you're not?) Jacques Derrida* argued that meaning is never immediate or fixed. In his concept of *différance* (a deliberate

* Hannah Arendt (1906–1975) fled Nazi Germany, coined '*the banality of evil*', and lost half her friends by saying what she thought. She also had an affair with her married philosophy professor, Martin Heidegger, who later joined the Nazi Party. She eventually forgave him but never let anyone else off so lightly.

** Theodor Adorno (1903–1969) was a German philosopher, sociologist, and one of the 20th century's more joyless critics of modern life. A frustrated musician, he argued that mass culture, especially jazz (and he really had it in for jazz) dulled critical thought and encouraged passive conformity. A brilliant mind, maybe, but not someone you'd want to get stuck with at a party.

* Jacques Derrida (1930–2004) was a French philosopher who coined *différance* to argue that meaning is never fixed, always postponed. He was also once scouted by a professional football team but turned down the pitch in favour of poststructuralism. Which, depending on your feelings about French theory, was either a loss to football or a greater loss to clarity. Viva la *différance*.

misspelling of difference), he suggested that meaning arises not only through contrast (difference) but also through deferral: the idea that meaning is always postponed, shaped by what comes before and after. We understand a word or idea not in isolation, but in relation to others, over time. Generative AI, by contrast, bypasses this process. It delivers answers without ambiguity or hesitation.

In this kind of system, less mainstream or original ideas may gradually disappear from view, reducing the diversity of thought in public conversations and making it harder for fresh perspectives to gain ground³⁹.

Disinformation and the end of shared reality

AI will also accelerate the creation of ‘*deep fakes*’ and disinformation, making it even harder to trust what we see and hear online. In a world where any video or audio can be fabricated to look real, the very notion of evidence can be put into question. Bad actors, whether state-sponsored or independent, can use AI to generate convincing falsehoods at scale. This could further pollute the information ecosystem and make people cynical about the possibility of knowing anything for sure (which itself is a tactic to sow confusion).

Another societal concern is the potential erosion of public knowledge; the shared set of facts and understandings that allow a society to have meaningful debates. If everyone is receiving personalised, AI-curated information streams, do we lose the common ground necessary for meaningful dialogue?

We risk a scenario where there is no broadly accepted baseline of truth, only myriad micro-realities, each shored up by its own AI-provided ‘*facts*’. In a sense, the ‘*death of knowledge*’ could manifest as the death of shared knowledge – essentially an *epistemic collapse*, where the very basis of a common reality or a shared understanding of the truth falls apart. We may all become ‘*experts*’ in our own artificially confined realms, with little overlap or mutual comprehension. Democratic governance becomes very difficult under those conditions.

Governance and democratic risk

If AI systems guide not just what we watch and buy, but who to date or vote for, what happens to individual and collective self-determination? We could slide into a passive mode of existence, where we assume “*the computer knows best*”. Political philosophers have warned of ‘*algorithmic governance*’, where citizens defer to AI-driven systems for decisions that used to involve human judgement, from parole rulings to welfare allocations. While AI can certainly inform policy with data, letting it dictate policy is another matter.

Donna Haraway's *Cyborg Manifesto* imagined a world where boundaries between human, animal, and machine blur, not as a dystopia, but as a challenge to rigid categories of identity, power, and knowledge. Written in the 1980s, it now reads less like a metaphor and more like a manual. We are already partial cyborgs, outsourcing memory, calculation, even conversation to machines. But the deeper question is not whether technology augments us, but how it redefines us. If the human mind is no longer the sole site of knowing, then knowledge itself is no longer confined to biology. As Haraway suggested, the cyborg isn't just a fusion of flesh and tech, it's a political figure, one that forces us to confront who holds power when information and identity are technologically mediated⁴⁰.



Already, some news organisations – and indeed some AI consultants – use AI to write or co-write articles, and some governments use AI to monitor citizens. The balance of power could tilt further if we allow AI systems (which are ultimately tools of whoever controls them) to mediate all information and decision flows.

Such support would be most welcome where decision-making skills are already weak or compromised. **The risk isn't a sci-fi scenario of AI becoming self-aware and taking over; it's the more subtle risk of people becoming overly deferential to AI outputs and the companies behind them,** losing the drive or ability to question and participate actively.

On the other hand, AI offers some positive possibilities in governance. These systems can analyse complex datasets far more quickly than any human, helping policymakers spot patterns or inefficiencies, for example, flagging tax fraud or pinpointing public health trends that might otherwise go unnoticed. AI models can also simulate the potential impacts of policies or detect early signs of problems, potentially making public services more evidence-based and responsive. When used responsibly, such tools could augment decision-making and allow governments to act more proactively in the public interest.

Governance, then, needs more than good intentions. It needs robust regulation. Platforms steered by people like Elon Musk shouldn't get to define or rewrite the 'truth' for commercial or ideological convenience. Some AI firms already act as though copyright laws don't apply. When the U.S. Copyright Office published a detailed critique of AI firms' reliance on fair use, Shira Perlmutter, the Register of Copyrights, was promptly dismissed by President Trump, widely interpreted as retaliation for standing up to the industry⁴¹.

But that regulation must be realistic. Nations or blocs like the EU or UK that clamp down too hard will fall further behind. China is surging ahead under state control. The U.S. is rolling back safeguards. On 23 July 2025, Trump unveiled *'Winning the AI Race'*, centred on deregulation, infrastructure expansion, and export—a direct play for AI dominance⁴².

That plan rescinded Biden-era safeguards via Executive Order 14179 (*'Removing Barriers to American Leadership in Artificial Intelligence'*), pre-empted state rules and signalled that strict copyright enforcement or patchwork local laws would jeopardise U.S. competitiveness, especially compared with China. As Trump put it, *"You just can't do it... China's not doing it"*⁴³.

So yes, the UK, the EU, and others should act decisively, but they must avoid getting entangled in bureaucratic red tape and overly restrictive rules while others race ahead. This has become a geopolitical contest over who controls how knowledge is created, shared, and trusted.

That said, there is real potential in AI for public-interest applications. These systems process massive public datasets far faster than any team of humans, spotting tax fraud, modelling policy outcomes or flagging health trends. Deployed carefully, AI could make governance responsive and far less reliant on guesswork.

But that only works if design is strategic, intentional and inclusive. That means:

- Limiting platform power and preserving competitive structures
- Educating citizens and public-sector staff in AI literacy to spot nonsense
- Creating institutions with real authority, not advisory bodies, for oversight of AI's role

The EU AI Act and voluntary codes of ethics within the industry are early moves in that direction. Whether they succeed depends on enforcement, coordination and public awareness. Easier said than done, of course.

And perhaps no institution feels this tension more directly than in education, where society decides what's worth knowing and how we learn to know it.

10. Educational Perspectives: Learning, teaching, & the future of knowing

Education is at the forefront of the AI revolution. Classrooms and campuses are already grappling with how students and teachers should use tools like ChatGPT. On the optimistic side, AI tutors and adaptive learning systems promise to personalise education. A well-designed AI tutor can, in theory, adjust to each student's pace, give extra practice on concepts they struggle with, and provide instant feedback on exercises. Such tools could help close achievement gaps by providing every student with something akin to one-on-one tuition.

AI's greatest value in the classroom may lie in making support more inclusive, not by replacing teachers, but by helping them focus on deeper learning. There are even AI systems that assist students with special needs by providing tailored support (e.g. converting text to speech for dyslexic students or generating practice problems for those who need them). This could make education more effective and inclusive if implemented well.

From cheating risk to curriculum rethink

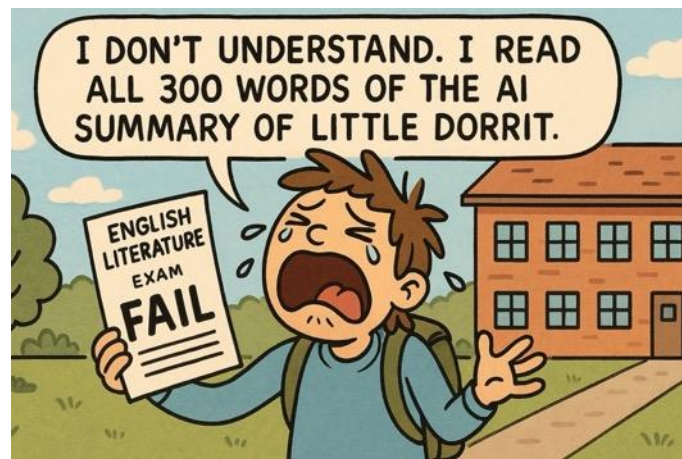
However, educators are understandably cautious, even alarmed. One immediate concern is academic honesty: if AI can do the work, will students bother to learn how to do it themselves? We've seen students use AI to generate essays, solve mathematical problems, and even answer exam questions. This isn't fundamentally different from past temptations (like copying from the web), but the scale and ease are new.

The risk is not just plagiarism, it's passive learning. If students rely too heavily on AI to produce answers, they miss the mental struggle that deepens understanding. If an AI always gives you a neatly packaged answer, you might not develop the grit to work through challenging problems or the habit of questioning and verifying information⁴⁴. Some early research suggests that when students use AI to help with homework, their independent problem-solving skills can suffer⁴⁵.

Whenever this paper seemed finished, another development arrived demanding inclusion. At the eleventh hour, OpenAI introduced *Study and Learn* mode in ChatGPT — a feature designed to guide students through the learning process using a form of the Socratic method. Rather than providing instant answers, it nudges users with questions, builds their understanding step by step, and checks comprehension along the way. A promising and well-intentioned development, though early signs suggest, predictably, it's being used more by anxious parents and teachers than by the students themselves*.

* As Paul Valéry said, "A work is never finished, only abandoned." At some point, one must stop tweaking, adding and editing, or risk entering a recursive loop from which no human or chatbot could emerge with their sanity intact. "Pens down" as they say in the exam halls, and increasingly in the boardrooms.

As AI tools are increasingly used to summarise long business reports and academic papers, we risk mistaking surface for depth. In some cases, marketers are producing content with AI, only for it to be summarised by the recipients using AI. The full document is read by neither the 'author' nor the 'reader'. As Paul Roetzer* puts it, relying on AI summaries is like trying to pass an English literature exam by only reading the CliffsNotes. You might grasp the plot, but you'll miss the language, nuance, and complexity.



If AI handles recall and routine tasks, curricula must shift towards interpretation, judgement, and problem framing. This means teaching students how AI works, where it goes wrong, when to use it (and when not to), how to critically evaluate its answers, and the ethical implications of AI. Crucially, they need to learn how to think for themselves.

AI literacy could become as essential as reading and maths.

Rethinking literacy, preserving agency

Rather than prohibiting the use of AI, educators could task students with using an AI tool to draft an essay and then critically revise it, learning to see where the AI might be wrong or simplistic. Or a class might collectively fact-check and improve a piece of AI-generated content, which teaches research and editing skills. The goal would be to harness AI's capabilities (speed, breadth, etc.) while still cultivating the uniquely human capacities that students need.

Some universities have issued guidelines encouraging transparent use of AI, e.g. if you use AI to help generate ideas or text, you must disclose it and reflect on what the AI did versus what you did. This way, students learn to use AI responsibly, a bit like learning to drive with supervision.

* Paul Roetzer is the founder of the Marketing AI Institute and presenter of the highly recommended podcast, *The Artificial Intelligence Show*. His co-presenter is the rather wonderfully named, Mike Kaput.

In an AI-rich classroom, teachers are not diminished; they're more vital. They guide students to challenge AI outputs, deepen their thinking, and keep learning rooted in human values. A good teacher can turn the AI into a tool for exploration while also warning students of its pitfalls.

Just as we teach arithmetic before introducing calculators, and source evaluation alongside internet search, we now need to teach knowledge literacy: how to judge AI-generated answers and know when deeper thinking is needed.

If AI is integrated thoughtfully, students can learn faster and think more deeply. If not, we risk raising a generation of button-pushers – confident with prompts, but shallow in understanding.

Which brings us full circle to the fear of knowledge's slow death.

Educators must grapple with fundamental questions: What skills remain essentially human? How do we assess learning when answers are a prompt away? How do we nurture *intrinsic motivation* when effort can be outsourced? What does it even mean to know something anymore? The answers will shape not just educational policy but the cognitive capacities of future generations.

At some point, though, we have to step back and ask a deeper question. If AI can answer anything, what does it really mean to know?



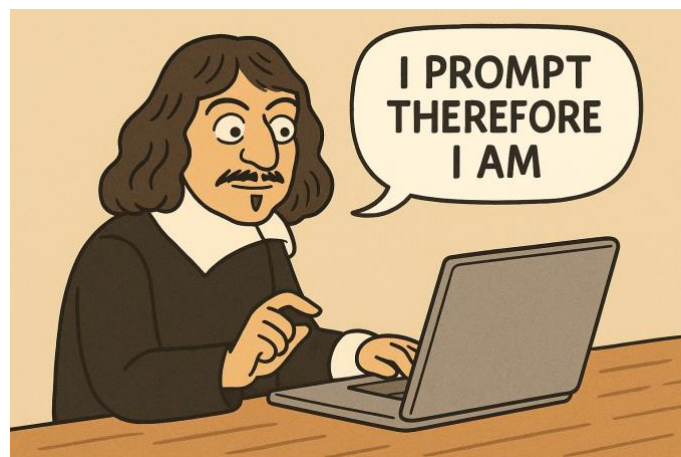
11. Philosophical Reflections: What does it mean to know in an AI world?

What does it mean to ‘*know*’ something when a machine can generate fluent answers on demand? Is knowledge just the possession of information, or does it involve understanding, judgement, and the ability to act with purpose? When AI can confidently respond to almost any prompt, the bar for ‘*knowing*’ doesn’t disappear, but it does shift.

Some argue that human knowledge is already changing. Memorising facts matters less. Knowing which questions to ask, how to interpret AI’s output, and when to trust it matters more. Expertise may no longer mean holding information in your head, but being able to frame a problem, spot a weak argument, and apply judgement in context.

In a world where answers are cheap, the skill is not in remembering, it’s in thinking. The real test may be knowing why you’re asking a question in the first place.

René Descartes* famously said, “*I think, therefore I am.*” He believed that doubt itself was proof of existence. If you can question, then you must exist. But what do we do with an AI that mimics doubt, writes passable essays on free will, and sounds like it knows what it’s talking about without understanding a word of it? The machine isn’t thinking. It’s predicting. But the output looks like thought. And that’s the unsettling bit. We’re no longer just asking, “*What do I know?*”. We’re asking, “*Who, or what, is doing the knowing?*”.



Of course, the machine isn’t actually reasoning. It’s not weighing ideas or doubting anything; it’s just laying down likely words in a plausible order. That’s the boundary between reasoning

* René Descartes (1596–1650) regularly withdrew to a warm room to escape the Bavarian winter and doubt everything he could. He was later tempted out by a generous offer from the Queen of Sweden to teach her philosophy at 5am each morning during the Swedish winter. He promptly caught pneumonia and died shortly after. Perhaps he should have thought a little more so he could have ached for longer.

Poor René had to fight to keep his place in this paper in the face of opposition from editors, but he deserves recognition, if only for coming up with one of the most famous lines in philosophy.

and simulation blurring. Not because the machine is thinking like us, but because its answers often sound like the sort of thing a thoughtful person might say.

If we become accustomed to easy answers on tap, will we still undertake the hard work of pondering difficult questions? Will we still stumble upon unexpected insights during a long, deep conversation or solitary study session?

Outsourcing thought: from Lyotard to Nietzsche

Philosophers have long worried about what happens when knowledge gets outsourced. In 1979, Jean-François Lyotard warned that as knowledge became externalised into databases and systems, the link between knower and knowledge would fray⁴⁶. He predicted a shift from understanding to utility: what matters is what can be stored, retrieved, and used. Anything not machine-readable might be ignored or discarded.

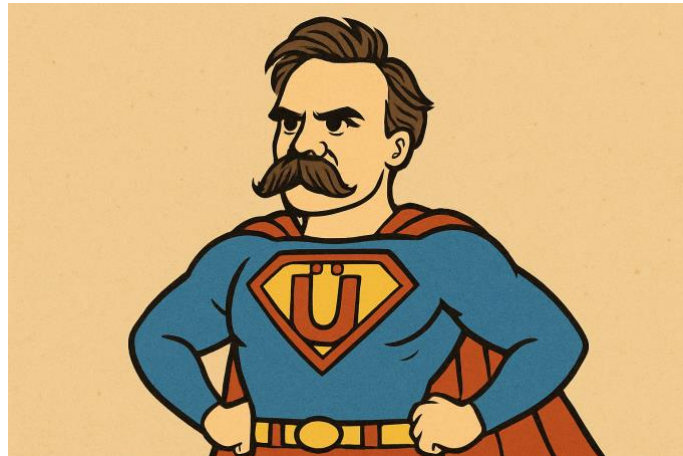
That sounds like now. AI is accelerating the very trends Lyotard saw coming. What society ‘*knows*’, from science to law to culture, increasingly lives in digital systems, accessed with no human in the loop. The benefit is access. The risk, Lyotard would argue, is the loss of meaning. **When knowledge is stripped of the context and human stories that give it depth, we risk becoming information-rich but understanding-poor.**

If machines remember and repeat, what’s left for us? Perhaps our job is no longer to store knowledge, but to make sense of it: to interpret, connect, and judge. Not memory banks, but meaning-makers.

Nietzsche* offered the image of the *Übermensch* - one who creates new values in the absence of certainties. We’ve built machines that can summarise Nietzsche, simulate wisdom, and recommend a mindful podcast to calm you down afterwards. But maybe that’s the point: if AI handles the trivia, it frees us to do the hard bit, thinking for ourselves. Nietzsche didn’t say we had to do it alone. Just that we had to do it on purpose⁴⁷.

There is another possibility worth considering. Some speculate that true superintelligence won’t replace us, but merge with us through brain-machine interfaces or neural implants. In this view, knowing becomes hybrid: no longer human or artificial, but co-generated.

* Friedrich Nietzsche (1844–1900), the owner of an impressively extravagant moustache, began as a devoted admirer of composer Richard Wagner, seeing him as a cultural revolutionary. The admiration didn’t last. As Wagner drifted into nationalism, Christianity, and antisemitism, Nietzsche rejected him, accusing the Ringmaster of pandering to the mob and dressing up prejudice as art. Their philosophical split was long, public, and conducted via increasingly pointed books. Nietzsche had a nervous breakdown after intervening to defend a horse from a vicious beating by his owner and was never the same again. Overall, he sounds like a decent bloke who just needed better PR.



That raises new questions: If your thought is created in partnership with a machine, does it still count as yours? If knowledge is shared across biology and silicon, who or what is doing the knowing? The conclusions we've drawn about '*human habits of mind*' being at risk may need rethinking. Perhaps those habits won't disappear but evolve. Not the death of knowledge, but its mutation.

How do we know when something knows?

Alan Turing's* famous test didn't ask whether a machine understands anything, only whether it can imitate a human well enough to fool another one in conversation. If it can, Turing suggested, maybe that's good enough. Intelligence by illusion.

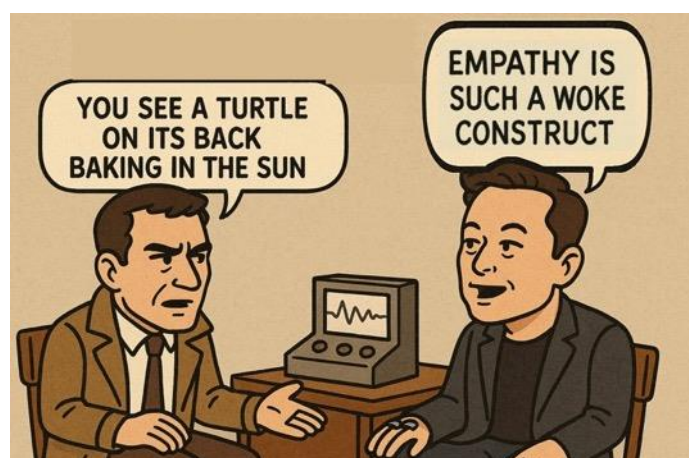
But that bar feels low now. Large language models routinely pass this kind of test without knowing what they're saying. They generate confidence, not comprehension. If the only test is "*do they sound smart?*", then yes, we're already living in the machine age.

Fiction took this further and darker. In '*Do Androids Dream of Electric Sheep?*', Philip K. Dick** introduced the Voight-Kampff test not for intelligence, but for empathy. Could a replicant feel? Or would it fake it too smoothly, missing the messy responses of real emotion?

* Alan Turing (1912–1954) proposed the test in his 1950 paper '*Computing Machinery and Intelligence*', suggesting that if a machine could hold a conversation indistinguishable from a human's, we might as well call it intelligent. Turing, being both brilliant and gay, broke codes that helped defeat the Nazis and was then prosecuted by the British government for gross indecency. They repaid his service with chemical castration. He died shortly after in 1954. A statue now stands in his honour at Bletchley Park, home of the British codebreaking service, which is nice, but the man himself might have preferred basic decency at the time.

** Philip K. Dick (1928–1982), was an American science fiction writer best known for raising unsettling questions about reality, identity, and what counts as human. His 1968 novel '*Do Androids Dream of Electric Sheep?*' (loosely adapted into the 1982 film '*Blade Runner*') introduced the Voight-Kampff test - an empathy detector used to spot replicants faking humanity. Late in life, Dick believed he'd been contacted by a divine intelligence via a pink beam of light and phoned the FBI to report it. When asked if he took amphetamines, he replied, "Not today."

The difference between these two tests is telling. The Turing Test checks for plausible fluency. The Voight-Kampff probes for moral substance. One asks: can it sound human? The other: can it care like one?



Maybe we'll never find a perfect test that can detect not just surface-level performance, but the capacity for doubt, for change, for moral reasoning. Something closer to judgement than output. But if such a test remains elusive, the burden falls back on us: to interpret, to decide, and to question what we're being shown.

Because until then, what we're really testing isn't machine intelligence. It's our own willingness to mistake a performance for the real thing.

Knowing how, not just knowing that

In the face of all this, what's needed may not be a better test, but a better kind of knowing — what we might call *meta-knowledge*. Knowing how AI works. When to trust it. When not to. How its answers are generated, and what its blind spots might be. Just as knowing how to find a book became essential once knowledge moved into libraries, and smart searching became essential when knowledge moved online, we now need to know how to work with AI, not just take its answers at face value.

Take medicine. A doctor in the near future may routinely receive AI-generated diagnoses. But her expertise will lie in knowing how to weigh that output, spot the missing context, and bring judgement, intuition, and care to the decision. That's not going away.

The case for wisdom

This shift only makes wisdom more important. Wisdom sits above knowledge as a blend of judgement, ethics, and knowing what truly matters. The more AI floods us with options and information, the more we need wisdom to decide what to do with it.

Which brings us right back to our old friend, Socrates, the philosopher who refused to write anything down. He only accepted the Oracle's verdict, that he was the wisest man in Athens*, once he realised it was a backhanded compliment. He was wise only in knowing how little he knew. Unlike his fellow citizens, who were generally untroubled by imposter syndrome, Socrates understood that genuine insight starts with humility and continues through questioning.

That questioning became the Socratic method: a systematic way of probing assumptions, exposing contradictions, and inching toward clarity.

Sound familiar?

That's not unlike iterative prompting, when it's done well. Prompting is a form of inquiry. It forces us to frame questions carefully, evaluate answers critically, and pay attention to what's missing as much as what's said.

Seen that way, the Socratic method isn't just a historical curiosity; it's a reminder that wisdom begins not with answers, but with better questions. ***Used well, AI informs, challenges, and extends our thinking. But it doesn't replace it.***

And that shift, from hoarding knowledge to navigating it, may be the most important move of all.

Knowledge is not dying.

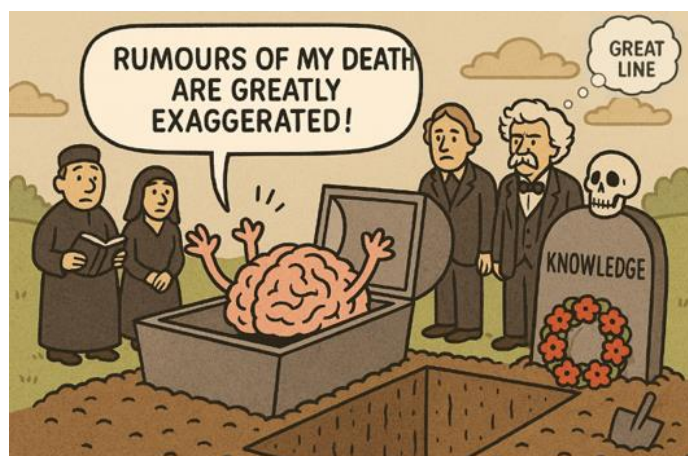
It's evolving.

Let's pull the evidence together.

* Women were unsurprisingly excluded from this unofficial competition. Socrates would probably not have approved. Although not a card-carrying feminist, he was unusually respectful of the wisdom of women. At least some of the time.

12. Conclusion

So, is this the death of knowledge? Not quite. But it is a profound transformation. AI is undeniably changing what it means to know, to learn, and to be an expert. It offers extraordinary opportunities to democratise access to information, accelerate discoveries, and augment human capabilities. At the same time, it threatens to erode memory, critical thinking, expertise, and the shared foundations of public knowledge if we handle it poorly.



AI is already delivering tangible benefits — from personalised tutoring that improves and extends attainment, to diagnostic tools extending medical expertise to under-resourced communities, and even the discovery of new antibiotics and therapies.

Our challenge is not to resist this change, but to shape it to serve human ends. That means investing in AI literacy, critical thinking, and the skills needed to thrive alongside AI^{48 49}. It means equitable access to AI tools and addressing the divides – digital, economic, educational – that could leave many people behind in an AI-driven world. It means recognising the limits of AI, valuing human judgement, and keeping humans in the loop for important decisions.

Above all, it means remembering that knowledge is not just information but a living practice; a way of engaging with the world, with others, and with ourselves. AI is a powerful tool, but it cannot replace the uniquely human capacity to care, question, be curious, and create meaning*.

The future of knowledge is not predetermined. It is up to each of us, as individuals, professionals, and citizens, to shape how AI is used, questioned, and integrated into our lives. I am optimistic. Not because AI is flawless; far from it. But because humans still have choices.

* These characteristics might not be uniquely human forever, but they still feel pretty central to what makes us, us. They're not sacred, but they are how we do thinking properly: consciously, deliberately, and sometimes even insightfully. If AI ever picks them up, it'll be copying us, not the other way round. So we might want to hang onto them.

We've already seen educators adapt, professionals retrain, and creatives find new ways to think and work alongside machines. I've trained people who were sceptical of AI at first, only to become confident users with better judgement, faster output, and renewed curiosity. The possibilities are enormous.

In short: ignore the hype, but don't ignore the shift. Knowledge isn't dead, but it is changing. And whether it gets shallower or deeper, more democratic or more distorted, won't be decided by the machines. It will be decided by us.

If we take that responsibility seriously, we may need to ask a different question: not just how we preserve knowledge, but whether knowledge is still the right thing to prize.

This isn't a retreat from knowledge; it's a recognition that our greatest value may lie elsewhere: not in trying to out-compete machines, but in choosing wisely how to use them.

Professor Nicholas Maxwell* spent years arguing that modern society, and academia in particular, had made a category error. We've pursued knowledge for its own sake, when we should have been pursuing wisdom: the capacity to judge well, to live well, to realise what is of value in life for ourselves and others.



The ancient Greeks had a word for this kind of wisdom: '**Phronesis**'. Aristotle** distinguished it from both technical skill (*techne*) and theoretical knowledge (*episteme*). **Phronesis is practical wisdom, not just knowing what is true, but knowing what to do. It's the kind of insight that guides sound judgement, ethical action, and the ability to navigate messy, real-world situations**

* Nicholas Maxwell (1937–2025), was a philosopher of science and long-time advocate for wisdom-inquiry. Largely ignored by academia for decades, presumably for the crime of being right too early. He spent much of his career arguing that universities should prioritise wisdom over knowledge, an idea now starting to sound less eccentric and more like a strategic survival plan.

** A late appearance from the early philosopher Aristotle (384–322 BCE), student of Plato, tutor of Alexander the Great. Like an Ancient Greek Gary Lineker, he seems to have been hanging around the metaphorical goalmouth of this conclusion, ready to tap this *mot juste* into the back of the net.

where the rulebook doesn't help. And after all, AI claims to be artificial intelligence, not artificial knowledge, and certainly not artificial phronesis.

In an age where AI can summon libraries of facts in seconds, perhaps it's time to hand over the knowledge part and focus on what only humans can do.

That includes wisdom — in Aristotle's sense of phronesis, or practical judgement — but also conscience, creativity, empathy, nuance, nous, and the capacity for doubt. Things that don't always play well with keywords and confidence scores, but which matter if you're trying to live, not just compute.

Above all, it encompasses *humanity*, not just in the moral sense, but in the full-blooded, often contradictory, not easily mapped experience of being human.

As the Stoics* remind us, the world will always change rapidly and unpredictably. What matters is our response. Wisdom demands effort, discipline, attention and a human mind willing to ask not just what is true, but what is good.

That feels less like a human advantage and more like a human obligation.



* The Stoics are my philosophers of choice, yet they didn't get a look in here, so forgive this gratuitous last page reference. Look out for more from them in the forthcoming sequel, *The Resistible Rise of Phronesis*.

Recommendations

How do we turn these ideas into action? Below I've listed recommendations for individuals, businesses, government and educators, before concluding with some more general ideas for us all. This is by no means a definitive list. But I hope it is a helpful start.

For individuals

1. **Understand how AI really works:** Know that AI doesn't think, it predicts. It doesn't understand truth or meaning. Treat its answers as suggestions, not facts.
2. **Keep your thinking skills sharp:** Don't hand over your memory, problem-solving, or writing too quickly. Practise thinking things through without always asking AI.
3. **Use AI to support curiosity, not replace it:** Use AI to open new lines of thought, not to shut them down. Avoid letting algorithms decide what you should learn or care about.
4. **Always question AI outputs:** Just because something sounds confident doesn't mean it's right. You don't have to accept the first response. Make a habit of checking, especially when the answer *really* matters.
5. **Balance convenience with challenge:** use AI to start and accelerate thinking, not finish it.

For business

1. **Keep the human in the expertise loop:** Let AI do the grunt work. But pay for people who can sense what matters, make judgements, spot when the machine is wrong, and provide... well... humanity.
2. **Make thinking part of the workflow:** Don't build systems where people just rubber-stamp what AI produces. Encourage challenge, interrogation, second opinions, and good questions.
3. **Train everyone how to use AI wisely:** Don't assume digital natives are AI-savvy. Teach all staff when to use AI, how to spot weak answers, and when to push back.
4. **Avoid feeding on your own recycled data:** If you let AI train on AI-created material, quality and originality will slowly disappear. Keep real-world thinking in the loop.
5. **Know where your information comes from:** Audit your decisions. Can you explain the thinking behind them? Or is it just what the AI said?

For government

1. **Make algorithms open and accountable:** When AI is used in public services to target service provision, decisions must be explainable. People have the right to know why they were accepted or rejected, prioritised or pushed to the back of the queue.
2. **Invest in public AI alternatives:** Don't leave all the power to big tech. Support open tools and research that anyone can access and understand.
3. **Teach the public how to understand AI:** Create education programmes so people know how AI works, where it goes wrong, and how to make their own judgements.

4. **Stop AI from controlling what people see:** Regulate how platforms rank and filter content. Make sure we still get a range of views – not just what AI thinks we'll like.
5. **Create a new public role to protect knowledge:** We may need a new body, like an 'AI Information Office', to track how AI affects truth, trust, and what we believe to be real.

For educators

1. **Teach how AI works – and why it can't be trusted or used unthinkingly:** Help students understand where AI is helpful, and where it goes wrong. Make its limits part of the learning.
2. **Design tasks that go beyond surface answers:** Don't just ask for essays or correct answers. Ask for thinking. Get students to explain, question, compare, and improve what AI provides.
3. **Make AI skills part of being literate:** Teach students how to use AI well – but also how to spot nonsense, bias, or overconfidence in its replies.
4. **Keep room for human thought:** Create spaces for open-ended discussion, debate, exploration and uncertainty – things AI can't do well (yet).
5. **Update what we mean by learning:** Move away from "*remember this*". Focus on "can you make sense of this?" or "*can you judge what matters?*" as a future-proof skill.

Cross-Cutting Recommendations

1. **Understand what AI is, not just how it works:** As we integrate AI into our lives, we must reflect on its nature, not just as a tool, but as a new category of actor in our systems of knowing. Is it a machine? An interface? A mirror? Or something else entirely?
2. **Make AI literacy a public priority:** Understanding how AI works – and how it fails – should be treated as a civic skill, not a specialist one. We need nationwide investment in public education, from elementary school children to senior leaders.
3. **Don't mistake information for understanding:** Just because AI can supply facts doesn't mean users understand them. Build habits and systems that reward context, reasoning, and the ability to explain, not just getting the answer fast.
4. **Keep humans in the loop for the decisions that matter:** Whether it's a hiring choice, a medical recommendation, or a welfare application, AI should support decisions – not make them alone. Human judgement must remain central and accountable.
5. **Build systems that challenge, not just optimise:** If AI is trained to reinforce what's already popular, we risk intellectual and creative stagnation. Whether in media, policy, or education, we must design for diversity, disagreement, and surprise.
6. **Make space for thinking and reflection:** In a world of constant pings, uninterrupted thought is a rare commodity. Get away from the screen, get outside and do a bit of daydreaming.



Postscript

This paper was written with the help of AI, not, I'd argue, out of laziness or hypocrisy (though I might confess to wilful provocation and poetic irony). It began as an experiment with the tools now available to us all, fuelled by a growing curiosity (that word again) about philosophy. Whether that marks decline or progress depends not on the tools themselves, but on how we choose to use them.

If nothing else, it proves that even reflections on the death of knowledge can now be co-authored by something that doesn't know it's reflecting. Perhaps that's the future: machines doing the thinking, humans doing the footnotes (and the jokes!).

And if you've made it this far without asking ChatGPT to summarise it for you, congratulations. You are clearly still in possession of a fully functioning attention span, which is a kind of knowledge in itself.

Glossary of Terms

Algorithmic Literacy

The capacity to understand and critically evaluate how algorithms shape the information one receives. It includes awareness of bias, opacity, and how algorithmic curation impacts what becomes '*known*' or visible in a digital environment. Often cited as essential for informed citizenship in an AI-driven age.

Anthropomorphism (in AI)

The tendency to attribute human qualities, like understanding, intention, or consciousness, to non-human entities such as AI.

Automation Bias

The human tendency to trust decisions made by automated systems, even when they're flawed or incorrect.

Banality of Evil

Hannah Arendt's concept that great harm can result not from monstrous intentions, but from unthinking obedience or bureaucratic conformity.

Black Box (AI)

Describes systems (especially deep learning models) whose internal operations are opaque or not readily interpretable by humans. This raises concerns about accountability and trust in domains such as law, where the reasons behind decisions are crucial.

Category error

A logical mistake where something is treated as if it belongs in the wrong type of category, like asking what colour Tuesday is.

Cartesian Dualism

The philosophical view (from Descartes) that mind and body are separate substances.

Cognitive Offloading

The use of tools or external systems to perform mental functions once handled by the brain: e.g. using calculators, calendars, or search engines. While efficient, habitual offloading may reduce long-term cognitive capacity, especially for memory, reasoning, or decision-making.

Cognitive Atrophy

The gradual decline of mental faculties due to disuse, often accelerated by overreliance on external aids like AI.

Commodification of Knowledge

A theme from Lyotard and others, referring to how knowledge is increasingly valued based on its market or instrumental utility (e.g. knowledge that can be sold, stored, or processed by machines) rather than its intrinsic worth.

Dataism

A quasi-philosophical belief, associated with thinkers like Yuval Noah Harari, that all aspects of life can ultimately be reduced to data flows. The article implicitly pushes back against this reductionist view by emphasising meaning, wisdom, and human agency.

Deepfake

AI-generated or AI-manipulated media (video/audio) designed to convincingly mimic real people. Raises ethical and epistemological issues about the nature of evidence, truth, and trust in a post-authenticity age.

Digital Amnesia

The tendency to forget information because we assume it's stored elsewhere (e.g. online or in devices).

Digital Divide / AI Divide

The inequality between those with access to digital (and now AI) tools, skills, and infrastructure, and those without. The AI divide includes disparities in awareness, confidence, and capacity to benefit from AI systems.

Epistemic Collapse

The breakdown of a shared understanding of truth or reality, often due to fragmented or personalised information ecosystems. In AI contexts, may occur when people receive conflicting but plausible answers generated by different models trained on divergent data.

Epistemic Humility

The intellectual virtue of recognising the limits of one's knowledge. Contrasts sharply with AI outputs, which often appear confidently asserted regardless of reliability.

Epistemology

The branch of philosophy concerned with the nature, sources, and limits of knowledge. Questions include: What is knowledge? How do we justify belief? What distinguishes knowing from merely having information?

Explainability

A design goal in AI where the system can provide human-understandable reasons for its outputs. The lack of explainability in most large models makes it hard to assess when and why an AI is wrong, or how it arrived at a conclusion.

Filter Bubble

A situation in which algorithms selectively present information aligned with a user's existing views, thereby reinforcing their worldview and reducing exposure to diverse perspectives. Coined by Eli Pariser.

Generative AI

A class of AI models that create new content (text, images, code, etc.) based on patterns learned from data without truly understanding the meaning.

Hallucination (AI)

A failure mode in generative AI where the system produces false or fictional content with apparent confidence. Not a bug per se, but a feature of how these systems predict likely text without '*knowing*' if it's true.

Hermeneutics

The theory and method of interpretation, especially of texts and meaning. Philosophers like Gadamer and Ricoeur focus on how understanding is always interpretive, situated, and historical.

Info-glut

The state of being overwhelmed by the sheer volume of available information, often leading to paralysis, superficiality, or a loss of meaning. A concern when AI floods channels with unfiltered or repetitive content.

Intrinsic Motivation

The drive to engage in an activity because it is inherently interesting, enjoyable, or meaningful, rather than for external rewards or pressures. It is associated with autonomy, curiosity, and a sense of purpose.

Knowledge-as-Information vs. Knowledge-as-Understanding

A key distinction in the paper. Information refers to discrete facts; understanding implies contextual, integrated grasp of meaning, implications, and relevance. AI excels at the former but struggles with the latter.

Logos, Pathos, Ethos

Aristotelian rhetorical appeals: logos (logic), pathos (emotion), ethos (credibility).

Lyotard's Externalisation of Knowledge

From **The Postmodern Condition**, Lyotard argued that in post-industrial societies, knowledge becomes external to the knower, stored in machines, databases, and systems, rather than residing in lived human experience. AI accelerates this trend.

Meta-knowledge

Knowledge about knowledge: how knowledge is produced, interpreted, and applied. In the context of AI, it includes understanding how systems generate answers, their limitations, and when to question their output.

Metacognition

Awareness and understanding of one's own thought processes — '*thinking about thinking*'.

Model Collapse

A degenerative cycle in which generative AI systems trained on AI-generated content become increasingly biased, repetitive, or error-prone, leading to loss of novelty and diversity in outputs. Represents a kind of epistemic entropy.

Nietzsche's Übermensch

A philosophical figure representing one who transcends conventional values and creates meaning in a godless, uncertain world. Invoked metaphorically in the paper as a challenge: will humans rise to new cognitive agency, or delegate meaning-making to machines?

Ontological Status (of AI outputs)

Refers to the nature of existence or '*being*' of something. AI outputs *appear* intelligent, but their ontological status is unlike human thought they have no intentions, awareness, or grounding in lived experience.

Performative Knowledge

Where the appearance of knowing is prized over substance, e.g. fluent AI responses. Charlie Munger labelled it *Chauffeur Knowledge*.

Personalisation (in AI)

Tailoring information or experiences to the user based on their data. While useful, it can fragment public discourse and erode shared realities.

Phenomenology

A philosophical approach that focuses on lived, first-person experience.

Phronesis

A type of practical wisdom or good judgement in everyday action. In Aristotle's philosophy, it refers to the moral and intellectual virtue of knowing how to do the right thing, at the right time, for the right reasons. Distinct from theoretical knowledge (*episteme*) or technical skill (*techne*). Our hero word!

Postmodernism

A movement in philosophy and critical theory sceptical of grand narratives, fixed truths, or objective knowledge. Relevant to the article's concerns that AI fragments knowledge and reduces it to what's useful or performative rather than meaningful or true.

Post-Truth

A cultural condition where appeals to emotion and personal belief are more influential than objective facts.

Prompt Engineering

The skill of crafting inputs to get the most relevant or reliable outputs from an AI model. A practical literacy, but also a reflection of how much AI 'knowledge' depends on human framing.

Recursive Learning / Degeneration

When AI systems learn from outputs of other AIs rather than human-generated data, errors and biases can compound. Known as a recursive loop, or in technical terms, *model degeneration*.

Situated Knowledge

A term coined by Donna Haraway, referring to knowledge that reflects the specific position, context, and perspective of the knower. Haraway rejects the idea of objective, universal truth.

Socratic Method

A form of dialogue based on asking questions to stimulate critical thinking and illuminate ideas. Named after Socrates, who distrusted written texts and preferred oral exchange as a path to understanding.

Stoicism

An ancient Greek and Roman philosophy focused on reason, self-discipline, and accepting what cannot be controlled. The paper briefly uses Stoicism to frame the human response to technological upheaval: we cannot control the spread of AI, but we can control how we respond.

Thoughtlessness (Arendt)

A concept from Hannah Arendt, who warned that evil (or injustice) can arise from the failure to think, reflect, and judge for oneself. Invoked in the paper to describe the danger of over-reliance on AI systems.

Turing Test

A test of machine intelligence proposed by Alan Turing in 1950: if a machine can hold a conversation indistinguishable from a human's, it passes. Not a measure of understanding, just imitation. Now, regularly passed by systems that don't know anything.

Utilitarianism

A moral theory that says the best action is the one that produces the most overall happiness. "*The greatest good for the greatest number.*"

Voight-Kampff Test

A fictional test from '*Do Androids Dream of Electric Sheep?*', designed to detect whether a subject feels empathy or is just very good at faking it. Unlike the Turing Test, it doesn't measure fluency, but emotional depth. Machines still fail. Some humans might, too.

Whole Brain Emulation

A theoretical process of scanning and replicating the entire brain's structure and function in a computer, effectively creating a digital copy of a human mind.

Wisdom

Traditionally seen as the highest form of knowledge – involving ethical judgement, practical understanding, and the ability to apply knowledge in a humane and context-sensitive way. Contrasted with mere access to information or factual correctness

Footnotes and references

¹ Sara Merken, “New York lawyers sanctioned for using fake ChatGPT cases in legal brief,” Reuters, June 26, 2023. (An incident where attorneys were fined for relying on AI-generated fake citations, underscoring the difference between access to information and true knowledge.)

² Plato, Phaedrus, 274c-275b. See e.g. John Uebersax, “Thamus and Theuth - Phaedrus 274b-279b” (includes the quote: “they will appear to be omniscient and will generally know nothing...having the show of wisdom without the reality.”).

³ Ibid. (“an elixir not of memory, but of reminding”). The phrase is discussed in Antonio Martínez, “Plato rejects writing by the mouth of Socrates,” antiquitatem.com, April 24, 2014, and other analyses of Plato’s attitude toward writing.

⁴ TIL Socrates was worried that writing would ruin students’ memory – we only know this because Plato wrote it down. (Paraphrase of a popular observation.) See the Reddit thread by user Breeze_in_the_Trees summarising this irony. Also, Bill Kowalski, “Socrates on the Forgetfulness that Comes with Writing,” New Learning Online, n.d

⁵ Shani Ealey, “Literacy By Any Means Necessary: The History of Anti-Literacy Laws in the U.S.” Oakland Literacy Coalition, Jan 12, 2022. (Discusses laws against teaching enslaved people to read.) See also PEN America, “Banned in the USA: The Mounting Pressure to Censor,” Aug 7, 2024, on recent book bans.

⁶ Andy Lee Roth, “Big tech algorithms: The new gatekeepers,” Reynolds Journalism Institute, July 31, 2024 [\[1\]](#). (Explains that proprietary algorithms run by Google, Meta, etc. act as gatekeepers of news and information, effectively determining what content is widely seen or hidden.)

⁷ Paul Roetzer, The Artificial Intelligence Show, Episode 158, “ChatGPT Agent, Grok 4, Meta’s Superintelligence Labs + Windsurfing,” July 22, 2025.

⁸ Pew Research Center, “How religion may affect educational attainment,” April 2016. (Example of socio-economic and cultural factors shaping access to education.)

⁹ Chris McClean & Almin Surani, “AI Is Deepening the Digital Divide,” InformationWeek, Nov 30, 2023. (Noting that unequal access to AI and tech can widen knowledge gaps.)

¹⁰ Teemu Birkstedt et al., “AI governance: themes, knowledge gaps and future agendas,” Internet Research (Emerald), 2022. (Touches on unequal distribution of AI benefits.)

¹¹ Olga Akselrod, “How Artificial Intelligence Can Deepen Racial and Economic Inequities,” ACLU, July 13, 2021.

¹² Raluca Csernaton, “Can Democracy Survive the Disruptive Power of AI?” Carnegie Endowment, Dec 2024. (On the need for inclusive AI policies to prevent social disparities.)

¹³ Alice Gomstyn & Alexandra Jonker, “Democratizing AI,” IBM Blog, Nov 5, 2024. (On efforts to make AI tools widely accessible) and Walter Yoshida, “Embracing the future of AI in the classroom: AI literacy, prompt engineering, and critical thinking in modern education,” Int. J. of Ed. Tech. in Higher Ed., Feb 26, 2024.

¹⁴ Khan Academy, “Khanmigo is used in more than 380 school districts in the United States, as well as in India, Brazil, and the Philippines,” Khan Academy Press Center, June 2025

¹⁵ Chirag Bhardwaj, “AI in Action: 6 Business Case Studies... Across Industries,” Appinventiv, Sept 27, 2024. (Examples include AI in healthcare diagnostics.)

¹⁶ LeewayHertz, “AI in knowledge management: use cases, benefits and development,” Nov 11, 2024. (On AI helping small businesses manage knowledge like larger firms.)

¹⁷ Christian Schellewald et al., “Algorithmic Literacy and the Algorithmic Divide: A Survey in Germany,” Frontiers in Communication, 2024. (Survey data on AI awareness: 88% heard of AI, 35% low knowledge, 13.9% feel high impact.)

¹⁸ The Death of Expertise, Tom Nichols’ 2017 book *The Death of Expertise: The Campaign Against Established Knowledge and Why It Matters*.

¹⁹ Melodi Pulliam, “Review of *The Death of Expertise* by Tom Nichols,” Journal of Intellectual Freedom & Privacy, Apr 9, 2018. (Nichols argues the internet gives laypeople the illusion of knowledge equal to experts, eroding respect for expertise.)

²⁰ Lyotard, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Translated by Geoff Bennington and Brian Massumi, Manchester University Press, 1984. (Originally published, forgivably in the circumstances, in French, 1979.)

²¹ Matthew Fisher et al., “Searching for Explanations: How the Internet Inflates Estimates of Internal Knowledge,” Journal of Experimental Psychology: General, 2015. (Study introducing the “Google effect” – people overestimate their own knowledge after using search engines.) Also discussed in Ross Pomeroy, “Is AI eroding our critical thinking?,” Big Think, Jan 29, 2025, and first of all in Sparrow, Liu & Wegner (Science, 2011).

²² Michael Gerlich, “Balancing Excitement and Cognitive Costs: Trust in AI and the Erosion of Critical Thinking Through Cognitive Offloading,” 2024. (Research suggesting habitual AI use can impair critical cognition.)

²³ Emily Warrender, “What does it mean to ‘know’ something in the age of AI?” Open Access Government, June 25, 2025. (Discusses studies on AI’s effect on memory and thinking skills.)

- ²⁴ IE University Center for Health & Wellbeing, "AI's cognitive implications: the decline of our thinking skills?" Feb 26, 2025. (Reports that younger individuals show stronger reliance on AI and potentially lower independent problem-solving scores compared to older individuals.)
- ²⁵ Faria Sanin et al., "Impact of artificial intelligence on human laziness, privacy and decision-making in education," Humanities & Social Sciences Communications (Nature), 2023. (Study of students in Pakistan/China; provides the statistics on perceived laziness, privacy issues, loss of decision-making ability due to AI.)
- ²⁶ International Monetary Fund (Kristalina Georgieva), "AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity," Jan 14, 2024. (Warns that AI could worsen inequality and erode incentives for education if not managed.)
- ²⁷ BuiltIn, "The Future of AI: How AI Is Changing the World," 2025. (Cites survey data: ~33% of workers believe AI could do many of their tasks; 44% expect significant skill disruption by 2028.)
- ²⁸ Strubell et al., "Training a large deep-learning model produces 626,000 lb of CO₂, equal to the lifetime emissions of five cars," MIT Technology Review, 7 August 2020. Note It is likely LLMs have become much more efficient since 2020.
- ²⁹ A. D. Selbst et al., "Fairness and Abstraction in Sociotechnical Systems," FAT* Conference 2019. (On the opacity of algorithmic decision-making and challenges of explainability.) See also IBM, "Explainability" – on the need for AI systems to explain reasoning.
- ³⁰ Daniel C. Dennett, interview with Noema Magazine, "The Illusion of Understanding," March 2023. See also From Bacteria to Bach and Back: The Evolution of Minds (2017), where Dennett discusses the dangers of anthropomorphising cognitive systems and mistaking competence for comprehension.
- ³¹ "Le doute n'est pas un état bien agréable; mais l'assurance est un état ridicule." From Voltaire's letter to Frederick William II of Prussia, dated 28 November 1770.
- ³² Hao et al., "AI and Memory," Memory, Mind & Media (Cambridge Core), 2023. (Discusses concerns that generative AI can flood society with informational "noise," untethered from human context, leading to a loss of shared meaning.)
- ³³ Ahad Ha'am (Asher Ginzberg, 1856–1927), quoted in Sam Atkinson, "The Philosophy Book: Big Ideas Simply Explained", DK Publishing, 2011. An excellent, comprehensive introduction to philosophy, did a great deal to inspire and inform this paper.
- ³⁴ David Dunning and Justin Kruger, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments", Journal of Personality and Social Psychology, Vol. 77, No. 6, 1999, pp. 1121–1134. DOI: 10.1037/0022-3514.77.6.1121
- ³⁵ PwC Global, "AI Predictions 2023" (survey of business adoption rates of AI, including generative AI). Summarised by Bernard Marr, Forbes, Sep 17, 2024
- ³⁶ Dr Boris Altemeyer, "All in the mix: AI is about augmentation, not just automation", Information Age, 29 October 2018, <https://www.information-age.com/ai-augmentation-automation-11850/>. Thanks to Ronnie for that one.
- ³⁷ Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Viking Press, 1963). Arendt argued that great harm can arise not from monstrous intent, but from passive, unquestioning conformity a warning newly relevant in an age where decision-making is increasingly outsourced to opaque algorithms
- ³⁸ Alice Gomstyn & Alexandra Jonker, "What is model collapse?" IBM Blog, Oct 10, 2024. (Describes how generative AI models trained on AI-generated data can enter "model collapse," compounding errors and producing increasingly homogeneous outputs. Over successive generations, such models lose "long-tail" information, narrowing the breadth of knowledge and reinforcing common biases.)
- ³⁹ Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.
- ⁴⁰ Donna Haraway, *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*, in *Simians, Cyborgs and Women: The Reinvention of Nature* (New York: Routledge, 1991), pp. 149–181.
- ⁴¹ Shira Perlmutter was dismissed days after her office released a report critical of AI firms' fair-use claims. Her firing sparked allegations of retaliation against copyright enforcement. AP News, May 2025. The legality of the dismissal is contested.
- ⁴² On 23 July 2025, Trump unveiled the AI Action Plan aimed at deregulation, data-centre expansion and broad export of U.S. AI technology to maintain global leadership. The Verge, July 2025
- ⁴³ "Trump orders AI action plan and more work erasing Biden's AI efforts", Reuters, 23 July 2025
- ⁴⁴ Binny Jose et al., "The cognitive paradox of AI in education: between enhancement and erosion," Frontiers in Psychology, April 14, 2025. (On how over-reliance on AI tools may diminish students' motivation and critical engagement.)
- ⁴⁵ Daniel Dillu, "How Over-Reliance on AI Could Lead to Cognitive Atrophy," Medium, Oct 8, 2023. (Examples of students using AI for homework and the potential long-term effects on learning.)
- ⁴⁶ Jean-François Lyotard, *The Postmodern Condition: A Report on Knowledge*, 1979. (Lyotard wrote that knowledge in post-industrial society becomes "exteriorised" from knowers; e.g., "the traditional link between knowledge and the knower is broken" and knowledge that can't be translated into machine-readable form "will be abandoned.")
- ⁴⁷ Friedrich Nietzsche, *Thus Spoke Zarathustra*, trans. R. J. Hollingdale (London: Penguin Classics, 1961). Nietzsche's Übermensch was a response to nihilism, a vision of humanity transcending itself by creating new values in a godless world.
- ³⁶ Pia Malaney, "Breaking the Moat: DeepSeek and the Democratization of AI," INET Economics Blog, Feb 10, 2025. (Argues for investments in AI literacy and open access to AI to ensure broad benefits.)

⁴⁸ Sandra Wachter, “*Empowering Public with AI Literacy*,” Harvard Gazette, Oct 26, 2020. (On the importance of AI education for the general public to maintain autonomy)

⁴⁹ Michal Wieczorek, “*Why AI will not Democratize Education: A Critical Pragmatist Perspective*,” Philosophy & Technology, Apr 21, 2025. (Stresses the need for critical thinking skills and human-centric teaching alongside AI tools.)

Contact

Trevor Lambert | Glass Halo Ltd | Trevor.lambert@glasshalo.co.uk