White Paper

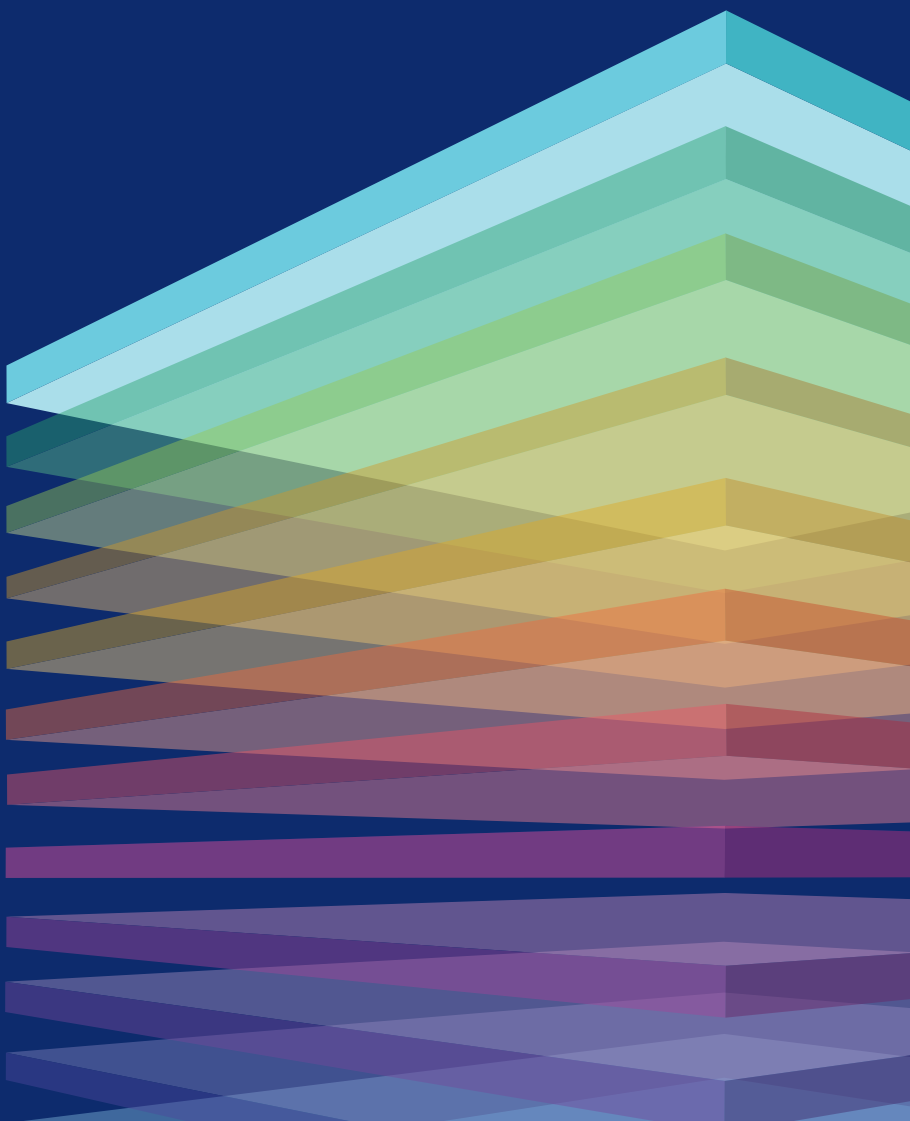
Sage

Human Skills for an AI Age

How today's business schools can develop tomorrow's leaders

Authored by:

Tom Chatfield



‘AI is reshaping how instructors and students approach education, and business schools can lead the conversation. This paper outlines critical elements for integrating AI into classroom learning. Discussion can’t just be around assignments and protocols of how students report their use of AI. Educators need to move towards helping students reflect on their own knowledge and understanding of the topic and how AI can support and enhance their learning. This paper is an important step in that direction.’

Professor Jeanine W. Turner, McDonough School of Business and Graduate School of Arts and Sciences, Georgetown University

‘New AI technologies led by generative AI will transform how businesses operate and how we work in and with them. Although we are only just beginning to understand the forms of this transformation, AI will augment and change rather than automate how we work. It will require us to develop new creative and critical thinking skills to continue to be at least equal partners with AI and to differentiate businesses from each other. This white paper sets out the agenda for business schools to deliver the AI+ business leaders of tomorrow.’

Professor Neil Maiden, Bayes Business School, City St George’s, University of London

‘Current conversations about AI are often polarized. AI is either portrayed as the savior of humankind, or as a guarantee for failure and demise. This white paper presents us with an integrative framework: toward a dual understanding of AI, rather than continuing the duel that is often propelled in public and scientific discourse. A better understanding of AI capabilities and, importantly, a focus on human capabilities such as critical thinking can be a fruitful angle to an AI-ready business education. Approaching a dual understanding of AI already in business schools is crucial for shaping a more nuanced and pragmatic way forward.’

Professor Christine Moser, School of Business & Economics, Vrije Universiteit Amsterdam and Associate Editor at Academy of Management Learning & Education

Introduction

Powerful, plentiful Artificial Intelligence (AI) is no longer a distant prospect. This poses a clear challenge for business schools: Equip students with the skills to lead in an AI-saturated world while upholding ethical standards and fostering personal and organizational development. Yet what it means to meet this challenge is uncertain. As AI accelerates and diversifies—most notably in the form of Generative AI, but also via other increasingly powerful and affordable applications of Machine Learning—its full potentials and risks remain unclear. Both in business schools and beyond, established approaches to learning, research and assessment are increasingly fragile. Nobody yet knows how AI will bring the greatest value—and where the boundaries lie between hype, disruption, and transformation. To help navigate these complexities, this paper presents *the DUAL framework*, focusing on four practical pillars for strategic resilience:

1

Demystify AI

Clarify the capabilities, limitations and nature of different systems to ensure students and faculty can use and experiment with them effectively.

2

Upskill critical thinking

Teach critical thinking about both individual cognition and its context in order to mitigate risks and maximize human potential, emphasizing the value of questioning assumptions, evaluating evidence and testing theories.

3

Augment human abilities

Ensure technology enhances rather than undermines human capabilities through a rigorous approach to data, ambiguity and oversight, designing learning experiences that leverage AIs' and humans' complementary strengths.

4

Lead through collaboration

Cultivate collaborative models of leadership, convening faculty, industry and students to explore the real-world contexts of AI's deployment—and the human needs it should serve.

This paper recognizes that business schools must work within the realities of accreditation pressures, resource constraints and stakeholder expectations. Adopting the DUAL framework in ways that feel achievable will help to develop learners, researchers, and leaders who are not only proficient in leveraging AI but also equipped with the critical, collaborative, and ethical skills needed to navigate an increasingly complex world with confidence. As the framework's four pillars emphasize, human insight and oversight are foundational activities, integral to the opportunities and responsibilities of the 21st century. The focus, therefore, is on the *dual* capabilities of humans and AI, working together and with humans in the driving seat, rather than a mutually diminishing *duel* between ourselves and machines.

Suggested citation: Chatfield, T (2025). *Human skills for an AI age: How today's business schools can develop tomorrow's leaders* (White Paper). London: Sage. doi: 10.4135/wp254070

The focus, therefore, is on the *dual* capabilities of humans and AI, working together and with humans in the driving seat, rather than a mutually diminishing *duel* between ourselves and machines.

Part 1: Demystify AI

Too often, debates around AI are dominated by hype, over-simplification or speculation, obscuring the practical and immediate challenges faced by researchers and educators. As the scholar of AI's social impacts Kate Crawford puts it in her book, *Atlas of AI*:

> *Artificial intelligence is both embodied and material, made from natural resources, fuel, human labor, infrastructures, logistics, histories, and classifications… In fact, artificial intelligence as we know it depends entirely on a much wider set of political and social structures.*[1]

Crawford's emphasis on technology's physical, particular existence is an important corrective to the assumption that AI is weightless and abstract. For all its power, modern AI is neither magical nor monolithic. Each system reflects specific design decisions, training data, and intentions; each deployment affects real lives, rights and the natural world. Used wisely, AI promises transformation and empowerment. But used uncritically, it can exacerbate dangers ranging from miscommunication and bias to exclusion, manipulative deception and environmental harm.[2] To demystify AI, it's thus vital to understand the technologies driving its current impact—and something of their history.

The historical context

While the term 'Artificial Intelligence' was coined in 1955[3], the field has evolved dramatically since then—and encompasses a wide range of capabilities and technologies. Early 'symbolic AI' focused on encoding human expertise into explicit rules, but proved too brittle for the complexities of real-world data. Today, the dominant force is 'statistical AI,' particularly **Machine Learning (ML),** where systems learn patterns from data without explicit programming. **Deep Learning,** a specialized form of ML using multi-layered artificial 'neural networks,' excels at handling complex patterns like those found in language and images. The most recent development, **Generative AI,** builds on Deep Learning to create novel content: text, images, audio, video, code. **Large Language Models (LLMs),** like ChatGPT and Gemini, are trained on vast datasets and can generate remarkably human-like outputs at superhuman speeds and scales. Such technologies are used, often in combination, to power countless applications.

- **Recommendation systems:** E-commerce giants like Amazon use ML to analyze customer behavior and tailor product recommendations, driving sales and enhancing customer experience.
- **Computer vision:** Retailers employ computer vision to monitor stock levels, optimize product placement and even analyze customer movement patterns within stores, leading to improved efficiency and sales.
- **Robotic process automation (RPA):** Insurance companies automate routine claims processing with software 'bots,' freeing human agents to handle more complex and nuanced cases.
- **Natural language processing (NLP):** Law firms leverage NLP, often powered by LLMs, to analyze legal documents, extracting key clauses and identifying potential liabilities more efficiently than traditional methods. However, it's important to note the limitations of such systems in understanding nuance, context and their tendency to 'hallucinate' plausible but false information.
- **AI agents:** These sophisticated systems combine multiple AI technologies, including Generative AI and LLMs, to understand instructions, interact with software and make decisions with increasing autonomy. They can potentially automate tasks ranging from market research and web design to coding, customer support and trading—although, thus far, they tend to be fragile in real-world situations.[4]

1 Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press, p.8.

2 See Ryseff, J., De Bruhl, B. & Newberry, S. J. (2024). *The root causes of failure for artificial intelligence projects and how they can succeed: Avoiding the anti-patterns of AI*. RAND Corporation. **www.rand.org/pubs/research_reports/RRA2680-1.html** and the classic paper on AI's wider societal impacts Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). *Association for Computing Machinery*, New York, NY, United States, 610–623. **https://doi.org/10.1145/3442188.3445922**

3 McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12. **https://doi.org/10.1609/aimag.v27i4.1904**

4 For a useful exploration of AI Agents' limits around coding tasks, see Husain, H., Flath, I. & Whitaker, J. (2025, January 8). *Thoughts on a month with Devin*. **www.answer.ai/posts/2025-01-08-devin.html**

As the above illustrates, it's unhelpful to make abstract pronouncements about such a diverse and dynamic range of tools. Future leaders must be able to cut through the bluster, identify appropriate use cases for different technologies, and make informed decisions about *when, how* and *why* to deploy them.

Lessons from the history of Information Technologies

The challenges posed by AI may be new, but many of the fundamental questions it raises are as old as humanity's relationship with information technology. As Neil D. Lawrence—Cambridge's inaugural DeepMind Professor of Machine Learning, and a former director of Deep Learning at Amazon—argues in his book, *The Atomic Human,* the history of information technology can be seen as a series of 'bottlenecks' being removed.[5] Key milestones include:

- **Written records:** Revolutionized law, commerce, culture, and governance by allowing records and ideas to be transmitted independently of human speech and memory.
- **Printing:** Enabled the mass dissemination of knowledge by unblocking the bottleneck of manual reproduction, sparking religious reformations and scientific revolutions.
- **Digital computing:** Ushered in an era of unprecedented connectivity by unblocking further bottlenecks of information-generation, sharing and processing.

AI represents a new phase in this progression. By automating pattern recognition, content creation and decision-making, it potentially removes the bottleneck of 'intelligent' analysis as a scarce resource. Tasks that once required human cognition—processing language and speech, object recognition, advanced problem-solving—can be performed at speeds and scales far beyond biology's limits. Yet it has also never been easier to move fast in a misguided direction, create a convincing fake, make misleading claims seem enticing, or abrogate responsibility to opaque systems.[6] As the philosopher and psychologist Alison Gopnik has argued, AIs are 'cultural technologies,' allowing us to interrogate information as never before—but also demanding newly sophisticated ethical and intellectual oversight.[7] From human rights and environmental impacts to legal accountability and the exploitation of intellectual property, the gamut of risks woven around AI demands a remarkable range and sophistication of human engagement. More, not less, is being asked from tomorrow's leaders.

New skills for abundant decision-making

Like the information revolutions that came before it, AI's unblocking of the bottleneck of human decisions is in effect a transition from scarcity to suffusion. Here is Ethan Mollick, Wharton Professor and AI expert, writing in November 2024:

> *We are just not used to abundant 'intelligence' (of a sort). Don't ask for an idea—ask for 30. Don't ask for advice—ask for many strategies. Pick and modify as needed.*[8]

The rise of digital photography offers a helpful analogy. With the removal of physical film's limitations, trillions of photographs are now taken and shared every day. While expertise and craft still differentiate professionals, the defining skills of the digital photography era have shifted towards selection, curation and contextualization, not to mention the use of digital images themselves as forms of data and insight. Similarly, the abundance of decision-making tools in an AI age puts a new emphasis on higher level human skills:

- **Framing problems:** Defining the scope and purpose of AI tools in specific contexts.
- **Evaluating outputs:** Critically assessing the quality, relevance, and implications of AI-generated suggestions.
- **Synthesizing insights:** Combining AI outputs with human creativity and domain expertise to produce actionable, accurate conclusions.

5 Lawrence, N. (2024). *The atomic human*. Penguin.

6 For some intriguing research into the potentially perilous persuasive powers of LLMs, see Luciano, F. (2024). Hypersuasion – on AI's persuasive power and how to deal with it. *Philosophy & Technology, 37*, 64. **https://doi.org/10.1007/s13347-024-00756-6**

7 Gopnik, A. (2022, July 13). Large language models as a cultural technology. **www.youtube.com/live/k7rPtFLH6yw**

8 Mollick, E. (2024, November). LinkedIn update. **www.linkedin.com/posts/emollick_we-are-just-not-used-to-abundant-intelligence-activity-7260303546749784064-1Sxb/**

- **Data literacy:** Appreciating how automated systems 'see' the world in terms of data—and what it means to source, create and interrogate this data effectively.
- **Systems and design thinking:** Understanding how parts of a complex whole interact, and what it means to design such systems to serve human needs.

This trajectory is being accelerated by unexpected players. Consider DeepSeek, an innovative Chinese AI startup founded in 2023, whose open-source R1 model gained attention in January 2025 thanks to performance rivaling that of giants like OpenAI and Google—at a fraction of the cost.[9] The emergence of systems like R1 suggests that the future of AI will be defined not just by scale and computational power, but by algorithmic efficiency and a widening of access driven by open-source models. Increasingly, competitive advantage may lie in the strategic application of AI, rather than ownership of the largest models—with expertise migrating from merely generating answers to asking the right questions and integrating insights into meaningful action.

Taking advantage of how AI actually works

At its core, the currently dominant approach to AI known as Machine Learning—and, in particular, the fields of Generative AI and Large Language Models (LLMs) within it—is a statistical engine, trained upon trillions of data points. When we 'speak' to an LLM, the experience resembles talking to a human being. This accessibility is transformative in its democratizing of activities like coding, content creation, and accessing information. But it can also be deeply misleading when it comes to AI's underlying nature and limitations. An LLM has no stable take on the world, no deep continuity between sessions, and no empirical knowledge of actuality beyond the data it was trained upon. It generates responses based on patterns in the data it has been trained on, predicting the most statistically likely outcome or sequence of words. Indeed, there remains substantial debate around whether the current generation of LLMs can be said to 'understand' anything—while their responses can rapidly become less accurate when dealing with relatively simple variations or additional details.[10]

Example: An AI analyzing Porter's Five Forces for the electric vehicle industry might confidently cite non-existent market research studies or invent specific-seeming competitor actions, because these align with patterns in its training data about how such analyses are typically structured.

Example: An AI tasked with conducting a break-even analysis using a company's sales data might generate plausible-looking but incorrect calculations by mixing historical and projected figures, or inventing data points to fill gaps—while appearing to show its work with precise decimal places that create a false sense of accuracy.

Yet the very same mechanism that leads to errors can also afford—if well-understood—impressive opportunities for enhancing creativity and productivity. When well-prompted, LLMs can rapidly generate novel ideas, simulate scenarios, suggests analyses and draft solutions in ways that challenge human assumptions and open up new possibilities.

Example: In marketing strategy, a well-prompted AI might accelerate the creative process by suggesting dozens of campaign ideas, combining patterns from successful past campaigns with novel twists that spark new insights—thus providing a starting point for humans to improve, extend, and challenge their own thinking.

9 See **https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf**

10 For example, see this influential October 2024 preprint from Apple machine learning research: Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S. & Farajtabar, M. (2024). GSM-symbolic: Understanding the limitations of mathematical reasoning in large language models. https://doi.org/10.48550/arXiv.2410.05229

Harnessing the particular potentials of Generative AI, then, means paying close attention to the entwined strengths and limitations of probabilistic interactions:

- When 'often brilliant, but occassionaly breathtakingly dumb' is acceptable, an LLM can function as a tireless and inhumanly rapid collaborator, offering on-tap insights and suggestions in almost any field.[11]
- From summaries and translations to schedules, transcripts, routine messages, processing unstructured data, prototyping software and interacting with data, this 'fast but fallible' assistance can vastly accelerate routine tasks and make new approaches possible at remarkable speed and scale.
- But when accuracy, quality and distinctiveness matter, it's vital to contextualize and verify outputs against real-world data, facts and needs.
- Generative AI can be unpredictable and fragile when faced with novel scenarios or edge cases. Critical evaluation and domain expertise remain (for now) the province of humans.
- The biases and limitations of training data may also be embedded in hard-to-detect ways within outputs.
- However, human learners' and researchers' insights into data biases and limitations can be a strength if an AI is actively used to explore and challenge these.

In particular, the rise of AI systems developed outside of Western contexts underscores the inherent link between technology and the values of its creators. As noted above, China's DeepSeek offers a powerful demonstration of technical prowess—but also reflects the political and social context of its origin, being subject to censorship aligned with Chinese government policies.[12] This highlights a crucial point: *All* AIs are the products of specific choices and priorities, none of which may be obvious to their users—let alone debatable by them.

Placing a premium on practical experimentation

As the Dutch computer scientist, Edsger Dijkstra put it four decades ago:

> *The question of whether machines can think… is about as relevant as the question of whether submarines can swim.*[13]

In the context of inherently opaque, unpredictable and rapidly evolving systems, what matters isn't so much speculation about what's 'inside' them as *investigating what practically they can (and cannot) do*—and how inputs and outputs can be assessed for accuracy, safety, and usefulness. This is an iterative, shared and reflective process. Consider the following set of guidelines, which sketches some of the norms a business school could embrace to help students and faculty engage practically with Generative AI's capacities:

- **Give permission and clarity:** Begin with clear guidance on when and where Generative AI can be used, how this usage is to be reported and reflected upon—and what a structure for iteration and experiment looks like. Emphasize practical, co-developed norms rather than exhaustive documentation.
- **Lean into condensing and explaining:** While final results need to be carefully verified, LLMs can rapidly and capably synthesize vast amounts of information, summarize diverse documents, and translate complex ideas into accessible formats, allowing learners individually and in teams to rapidly generate rich context for (say) business case scenarios or project work.
- **Prioritize prototyping and experimenting:** When well-prompted with the 'story' of a desired scenario, an LLM allows for rapid drafting, modelling and iteration on synthesized context, enabling users to test and refine detailed ideas and possibilities quickly. Integrate 'teaching the AI' into topics, classes and group work.
- **Embrace coding, models and personas:** Students can use LLMs to write code, create interactive models of business problems, simulate scenarios, draft mock proposals and interact with different personas within a scenario. Learning structures should also include individual and collective critical reflection, encompassing what it means to prompt an AI to achieve better results (and what 'better' means).
- **Integrate reflection and verification:** Generativity and efficiency come at the potential cost of opacity and unreliability. Users may not fully understand how outputs were generated or whether they are accurate, and thus need to build independent verification, reflection, reporting and testing into scenarios. This verification and reporting is a powerful learning opportunity.

11 See Gary Marcus' prescient essay. Marcus, D. (2022, December 1). *How come GPT can seem so brilliant one minute and so breathtakingly dumb the next?* **https://garymarcus.substack.com/p/how-come-gpt-can-seem-so-brilliant**

12 Lu, D. (2025, January 28). We tried out DeepSeek. It worked well, until we asked it about Tiananmen Square and Taiwan. *The Guardian*. **www.theguardian.com/technology/2025/jan/28/we-tried-out-deepseek-it-works-well-until-we-asked-it-about-tiananmen-square-and-taiwan**

13 Lecture delivered at the ACM 1984 South Central Regional Conference, November 16–18, Austin, Texas. **www.cs.utexas.edu/~EWD/transcriptions/EWD08xx/EWD898.html**

Beneath all this lies one of the deepest tensions within Generative AI. While it may appear to offer instant answers and automate complex tasks, it still requires considerable human expertise to ensure outputs are reliable—and to understand what it means to improve them in a useful direction. Moreover, without some capacity to check and reflect upon outputs, users risk mistaking confidence for accuracy, bias for insight, and noise for meaning. An LLM is not a search engine, and a probabilistic output needs to be carefully differentiated from an authoritative record.

This tension underlies the challenge for business schools. Using AI responsibly and effectively means assessing and contextualizing its outputs. Yet doing so necessarily depends upon the specific data, systems and risks involved in a system's deployment—and domain expertise that learners may well lack. Hence the significance of scenario-based learning, in-depth engagements with industry—and practical over theoretical skills, formed by seeing the outputs of actual systems put to the test. Here, educators have a key role to play by creating contexts for meaningful (and potentially highly engaging) experimentation, validation, and critique. As Kellogg School of Management professor, Robert L. Bray puts it in a January 2025 paper describing his use of LLMs in a data analytics class:

> *Let the AI be your muse. ChatGPT transformed my elective class from one struggling section to three brimming sections in just two years. This happened not so much because students like the technology but because I like the technology. Simply put, ChatGPT made investing in my class fun again. AI allowed me to do things that had never before been done in the classroom. I got hooked on finding AI-empowered teaching innovations.*[14]

The more that LLMs are treated as flawed-but-powerful tools to be playfully tested rather than magical boxes, the more likely it is that uses will be found and adopted.

“Let the AI be your muse. ChatGPT transformed my elective class from one struggling section to three brimming sections in just two years. This happened not so much because students like the technology but because I like the technology. Simply put, ChatGPT made investing in my class fun again. AI allowed me to do things that had never before been done in the classroom. I got hooked on finding AI-empowered teaching innovations.

– Robert Bray, Kellogg School of Management

14 Bray, R. (2025). *A tutorial on teaching data analytics with generative AI.* **www.kellogg.northwestern.edu/faculty/bray/doc/chatgpt/chatgpt.pdf**

Practical recommendations

1. Create clear codes of conduct for AI use and experimentation

- Establish clear, pragmatic codes of practice that encourage confident experimentation while ensuring accountability and ethical rigor.
- Co-design and regularly update these codes in the light of learning and discovery. Clarity and trust between students and faculty are vital, while fear and uncertainty are likely to breed secrecy and mistrust.
- Provide ethical guidance for addressing risks like bias, over-reliance, privacy, environmental impacts and intellectual property infringement, ensuring students and faculty can engage with AI confidently and responsibly.

2. Build deep AI literacy into learning

- Explore the specific mechanics of different AI systems, including Generative AIs' reliance on training data, human oversight and their potentials for hallucinations and bias.
- Use hands-on exercises to explore how LLMs' outputs vary based on prompts and assumptions, emphasizing critical evaluation rather than passive acceptance.
- Embed this literacy across disciplines, showing how AIs' applications and limitations apply differently in contexts like marketing, operations and strategy.

3. Turn generative AI's probabilistic nature into an opportunity

- Encourage students to learn by experimenting widely, sharing insights and regularly updating their assumptions as systems evolve. Educators should also be ready to experiment and model an exploratory, practical approach.
- Teach specific problem-framing skills to elicit diverse LLM outputs, followed by strategic engagement with these outputs via 'meta' skills like curation, interrogation and systems thinking.
- Train students to critically assess patterns and anomalies in AI outputs, combining them with human judgment to draw actionable insights.

Part 2: Upskill critical thinking

In an era where Generative AI systems can generate plausible responses to any query with remarkable speed and fluency, our natural cognitive tendencies can lead us into uncritical acceptance of information. As Daniel Kahneman, the Nobel laureate and pioneer in behavioral economics, observed:

> *This is the essence of intuitive heuristics: when faced with a difficult question, we often answer an easier one instead, usually without noticing the substitution.*[15]

When confronted with complex problems, people often default to fast, intuitive answers based on their current knowledge and feelings, without pausing to assess their validity. This is because human minds evolved to handle the social and physical challenges of a pre-technological world, rather than to critically evaluate the outputs of sophisticated language models. Three vulnerabilities deserve close attention:

- **First, we naturally equate fluency and familiarity with truth.** The more easily information flows, the more likely we are to accept it. This heuristic served us well when such factors correlated with safety and reliability, but makes us vulnerable to systems capable of generating endlessly plausible content with perfect confidence (not to mention human beings eager to exploit such systems).
- **Second, we are vulnerable to confirmation bias.** The tendency to seek out and embrace information that supports our existing beliefs while dismissing contradictory evidence is widespread. AI systems, being explicitly designed to be helpful and meet user expectations, can reinforce this tendency and echo human cognitive weaknesses. Similarly they often lean towards an 'average of the average' response, avoiding extremes and potential controversy even when these might be appropriate.
- **Third, we find it hard to recognize the limits of our knowledge.** When faced with difficult problems requiring careful analysis ('What are the long-term implications of this technology?'), we are often prepared to accept simple answers based on limited information and experience, even if this is manifestly inadequate in empirical terms ('Does this AI produce convincing answers fast?').

If a business student or leader asks an LLM for strategies to enter a new market without contextualizing their inquiry, the system may rapidly generate a compelling plan that overlooks potential risks or alternative perspectives. The AI's ability to produce coherent and persuasive narratives gives a false sense of confidence, flattering pre-existing prejudices and assumptions. In general, LLMs are adept at giving almost *any* idea a veneer of authority or coherence. This places an onus on human thinkers, leaders, and teams to adopt a 'testing' rather than 'confirmatory' mindset, and to put protocols in place to mitigate the overlaps between cognitive bias and technological possibility.

Practical techniques and ethical considerations

In the context of AI, critical thinking can be thought of as the process of contextualizing, analyzing, synthesizing and evaluating information—and of identifying and mitigating relevant systemic and cognitive limitations. In effect, critical thinking serves as a form of quality control for the human mind. It thus entails *pausing to seek cognitive reinforcements* in a variety of ways:

- **Question AI outputs:** Rather than accepting AI-generated outputs at face value, we critically assess their validity, relevance and potential biases, as well as using them to challenge and improve our own judgements. Among other things, this means using reliable external sources of data, independent human expertise, and prompting AIs via techniques that prioritize iteration, exploration and challenge.
- **Interrogate initial assumptions:** Rather than go with a single 'default' framing of any significant problem or question, we explore important issues from multiple perspectives, trying to unearth misleading assumptions and compare alternative priorities and lines of investigation.
- **Identify structural limitations:** We try to recognize what AI cannot 'see'—the nuances, contexts and implications that may not be captured in its training data—then mitigate this via human expertise and other systemic checks and balances. This in turn means drawing on others' research around training sets, biases, and the strengths and weaknesses of particular systems in particular contexts.

15 Kahneman, D. (2011). *Thinking, fast and slow*. Allen Lane, p. 12.

Ultimately, all data is the product of a process of measurement, processing or selection rather than a neutral resource. A defining human task in an information age is thus to move beyond the passive consumption of information towards active questioning of its provenance, significance and limitations. Key questions include:

- What does this information signify?
- How and why was it generated?
- How representative and reliable is it?
- What can and cannot be reasonably inferred upon its basis?
- How can this information best be tested and improved?
- What are its biases, exclusions and limitations?
- What else do we need to know to address the challenges we care about?

As researchers such as Timnit Gebru, Emily Bender, Abeba Birhane, Cathy O'Neil, Margaret Mitchell, Shannon Vallor and others have highlighted, structural biases, exclusions and distortions are embedded within much of the data many ML systems—and LLMs in particular—were trained upon.[16] It's thus essential for business schools to teach and address considerations including:

- **Diverse perspectives in AI development:** Emphasizing the importance of involving individuals from diverse backgrounds and fields in the design, development and deployment of systems. This helps to ensure that a range of perspectives and experiences are considered, reducing the risk of unintended negative consequences for particular groups.
- **Accessibility:** Considering the accessibility of AI tools and technologies for individuals with disabilities. Ensuring that AI-powered learning platforms and resources are designed to be inclusive and usable by all.
- **Leadership and followership:** Exploring the question of diverse leadership and followership in an AI age. Recognizing that diverse teams may interact with and be impacted by AI in different ways, and that leaders need to be sensitive to these differences.
- **Supporting underrepresented students:** Recognizing that students from underrepresented or less privileged backgrounds may face additional challenges in an AI-enabled world, such as a higher perceived risk of failure or fewer resources to invest in learning. Providing targeted support and mentorship to ensure all students have the opportunity to succeed.
- **Environmental and social responsibility:** Teaching and debating what it means to train and deploy AI ethically in this wider sense, looking at its impacts on areas like human rights, working conditions and the natural world—and the wider infrastructure of data centers, energy supply and human moderation that is often hidden from view. This also includes issues such as algorithmic discrimination, surveillance and the potential for AI to be used to manipulate or control individuals and communities.

There is no such thing as a perfectly impartial, representative or exhaustive dataset, and thus no such thing as a perfectly impartial, fair or insightful AI. No matter how much is known, the ethically and critically discerning use of AI in general—and LLMs in particular—will always require a reflective negotiation with its limitations and societal context.

No matter how much is known, the ethically and critically discerning use of AI in general —and LLMs in particular—will always require a reflective negotiation with its limitations and societal context.

16 Among other resources, see the work of Black in AI, founded in 2017 by the computer scientists Rediet Abebe and Timnit Gebru as 'a place for sharing ideas, fostering collaborations and discussing initiatives to increase the presence of Black people in the field of Artificial Intelligence'. https://blackinai.github.io/#/ - the work of Cathy O'Neil's algorithmic work consultancy OCAA - https://orcaarisk.com/ - and Shannon Vallor's 2024 book, *The AI mirror*. Oxford University Press.

Allocating time and attention wisely

It's worth emphasizing that critical thinking is inherently active and effortful. It *cannot* take place without the allocation of sufficient time and attention. These are often in short supply. From social media updates to overflowing inboxes and proliferating logins, both learners' and instructors' time can be a painfully scarce resource. This in turn means that a foundational challenge for business schools is to create contexts within which high quality self-reflection, debate, and collaboration can happen among students.

As Kahneman emphasizes in his work:

> *Cognition is embodied; you think with your body, not only with your brain.*[17]

The creation of places, spaces, and practices that support high-quality reflection is precisely about recognizing that humans are *not* machines; and that high-quality cognition is reliant upon factors such as rest, variety, time, focus, mutual support, and morale. As Microsoft's human factors lab noted in a study of online meeting patterns during the pandemic:

> *When participants had meditation breaks, brainwave patterns showed positive levels of frontal alpha asymmetry, which correlates to higher engagement during the meeting. Without breaks, the levels were negative…In sum, breaks are not only good for wellbeing, they also improve our ability to do our best work.*[18]

At their worst, the automated analysis and production of vast amounts of data risks disengaging, deskilling and overwhelming human workers and learners alike. The so-called 'techlash' against exploitative, demeaning or manipulative systems is only likely to grow in significance with the ubiquity of AI, as are both activism and legislation inspired by it.[19]

At their best, technologies like AI promise an increase in the quality of human time, attention, and engagement. But aligning them with this aspiration requires a wider engagement with the context of routines, spaces, and social norms: that which lies 'outside the box' in terms of permission and expectation. Opportunities to defend high quality time and attention include:

- Defining breaks, boundaries and protocols around communications timings and quantities, e.g. no email after 6pm or at weekends
- Setting up screen-free zones or moments within campuses and working days
- Regular reviews of time sinks or bottlenecks within software platforms or daily tasks, e.g. time-use reporting to spot issues with efficient use of VLEs
- An insistence that the discerning use of any technology entails saying 'no' as well as 'yes' to its offerings.

Productivity gains are rightly hailed as a key promise of AI. Without a complementary focus on the quality of time, self-reflection, dialogue, and active listening, however, such gains are all too likely to tip over into informational overload, distraction, and credulousness. As emphasized by an in-depth 2024 study of 666 learners of varied ages and backgrounds, uncritically 'offloading' cognition to AI tools is associated with reductions in both critical engagement and academic performance:

> *The findings revealed a significant negative correlation between frequent AI tool usage and critical thinking abilities, mediated by increased cognitive offloading… Furthermore, higher educational attainment was associated with better critical thinking skills, regardless of AI usage.*[20]

Unless embedded within a culture of critical engagement, AI may fulfill a fear about information technology ancient enough to have preoccupied Plato, offering its disciples 'not truth, but only the semblance of truth; they will be hearers of many things and will have learned nothing.'[21]

17 Kahneman, D (2011). *Thinking, fast and slow.* Farrar, Strauss and Giroux, pp. 50–51.

18 See *Research proves your brain needs breaks*. www.microsoft.com/en-us/worklab/work-trend-index/brain-research and associated research at **www.microsoft.com/en-us/research/project/the-new-future-of-work/publications/**

19 Consider the author and activist, Cory Doctorow's November 2022 coining of the term 'enshittification' to describe the degrading of digital services, their exploitation of users and what this signifies: *How monopoly enshittified Amazon* **https://doctorow.medium.com/how-monopoly-enshittified-amazon-83f42a585c3c**

20 Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), 6. **https://doi.org/10.3390/soc15010006**

21 Taken from from Benjamin Jowett's classic translation of *The Phaedrus*, which can be read for free online in full at http://classics.mit.edu/Plato/phaedrus.html. The full passage is a critique, delivered by the character of Socrates, of the risk that written words may encourage forgetfulness and ignorance in those reliant upon them (as opposed to the living practice of debate).

Critical thinking both as individual and social activity

While it can seem solitary, critical thinking needs to be seen in the context of business schools as *both* an individual skill and a collective practice. Individually, it involves self-reflection and the ability to analyze your cognitive strengths and limitations. Collectively, it thrives on diverse perspectives, project work, and collaborative research and problem-solving. In their 2024 paper, 'Critical Thinking in the Age of Generative AI', Barbara Z. Larson and her colleagues emphasize the importance of both dimensions:

> *Both individual and social critical thinking will be essential for students to succeed in the AI era. Students will need individual critical thinking to evaluate GenAI outputs, identifying poor-quality, inaccurate, and confabulated information… Social critical thinking will enable our students to identify missing perspectives, marginalized voices, and taken-for-granted societal assumptions in GenAI output.*[22]

In particular, individual critical thinking involves:

- Systematic evaluation of evidence and arguments
- Recognition of one's own biases and assumptions
- Active questioning of AI-generated outputs
- Metacognitive awareness of one's own thought processes

While social critical thinking encompasses:

- Collaborative evaluation of ideas and evidence
- Recognition of diverse perspectives and experiences
- Identification of systemic biases and cultural assumptions
- Collective resistance to groupthink and premature consensus

This once again highlights the double-edged nature of AI as a tool. When used uncritically, it can reinforce both individual cognitive biases and collective oversight. When used in a critically engaged and reflective context, however, precisely the opposite can be true. By actively interrogating and exploring the strengths and limitations of data, prompts, models and AI tools, the same technological context can become a spur to deeper critical engagement—and to investigations impossible without technology.

Example: By interacting with an AI model trained on financial reports, analyst calls, and media coverage from 2005-2008, business students can critically explore how institutional assumptions and market narratives evolved during the subprime mortgage crisis—revealing how cognitive biases and groupthink manifested in the language and logic of key decision-makers.

Below are some proven examples of techniques that can help individuals and groups think critically, including suggestions on how AI might play a part:

- **Premortem analysis:** A premortem is a strategy where a team imagines a project has failed, then works backward to imagine what might have caused that failure. This can help to identify potential pitfalls and challenges that might be overlooked due to overconfidence or confirmation bias. Such a framework also provides excellent opportunities for using AI to add richness, challenge and context, extrapolating from suggestions and helping find mitigations.[23]
- **Devil's advocacy/stress-testing:** Encouraging students to take on the role of a devil's advocate can similarly promote critical examination of ideas and plans. By deliberately challenging assumptions and proposals, students learn to identify weaknesses and consider alternative perspectives. Similarly, AI can be used to simulate different scenarios or generate counterpoints to a proposed strategy, helping to reveal potential flaws or areas for improvement.

22 Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K. & Colombo, L. A. (2024). Critical thinking in the age of generative AI. *Academy of Management Learning & Education*, *23*(3), 373–378. **https://doi.org/10.5465/amle.2024.0338**

23 For the origins of the premortem, see its inventor, Gary Klein, in the September 2007 issue of *Harvard Business Review.* **https://hbr.org/2007/09/performing-a-project-premortem**

- **Exploring structural biases:** Analyze AI outputs to identify and understand inherent biases–both socially, but also in terms of factors such as recency, confirmation, and availability—fostering a deeper awareness of how data and algorithms can influence outcomes. This process helps students develop a critical eye towards the information they receive, recognizing that even seemingly objective data can be shaped by structural distortions and societal prejudices.
- **Reframing key questions:** AI can actively be prompted to pose reflective questions or suggest alternative viewpoints, encouraging students to think more deeply about the issues at hand. This encourages metacognition, prompting students to examine their own thought processes and assumptions. It also emphasizes how AI can, when used well, become an active part of individuals' and teams' metacognitive apparatus.

It's important to note that all the above involve students (and educators) learning from one another, not just AIs. A significant opportunity for AI, moreover, is that it may be able to mediate such learning in a flexible, scalable, non-judgmental way. In cognitive terms, it may be easier to confess uncertainty and doubt to a machine than a tutor or colleague. High-performing teams and institutions should harness such dynamics to facilitate richer, franker, and more inclusive debates.[24]

Building critical thinking into business education and student assessment

To effectively upskill critical thinking, business schools should embrace the fact that the critically reflective oversight and management of AIs is likely to be one of *the* defining human skills of the next decade. Opportunities include:

- **Curriculum design:** Offer courses that combine business with philosophy, psychology, and ethics to provide a broader context for critical thinking. This exposes students to diverse perspectives and methodologies, enriching their analytical toolkit and enabling them to approach problems from multiple angles.
- **Incentivizing reflection and communication:** Assign work that requires students to reflect on their decision-making processes, the reasoning behind their conclusions and their use of technology. This encourages self-awareness and metacognition, prompting students to articulate their thought processes and identify areas for improvement.
- **Embracing orality and topical debate:** The emphasis on reflection and communication can also extend to structured debates on topical news items such as the recent Jaguar 'Copy Nothing' ad, encouraging students to articulate their positions–-with a particular emphasis on orally debating, exploring and defending arguments.
- **Faculty development:** Provide faculty with resources and training on how to teach critical thinking skills effectively, including strategies for fostering open dialogue and inquiry. This ensures that educators can promote critical thinking and provide students with the necessary support and guidance—not least in terms of psychological safety, active engagement, and tolerance.
- **Testing AI's limitations:** As this section has explored, AI models are trained on historical data, which may contain biases or lack representation of current trends and diverse perspectives. Discussions and assignments should explicitly aim to identify, test and mitigate this. Similarly, AI lacks comprehension in the human sense and cannot fully grasp nuances, emotions or ethical considerations—or factors such as causation—inherent to human thought and perception. This should be tested, debated and mitigated.

24 For an interesting perspective on attitudes towards AI facilitation in a therapeutic context, and reflections on where and how it may (and may not) be preferable to human facilitation, see Aktan, M. E., Turhan, Z & Dolu, I. (2022). Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Computers in Human Behaviour*, *133*, 107273. **https://doi.org/10.1016/j.chb.2022.107273**

Perhaps most imminently, the rise of Generative AI poses profound challenges to traditional assessment methods. If students can use AI to generate essays, reports, and code, how can educators ensure assessments are accurately measuring student learning and not simply the capabilities of AI? Addressing this requires a shift in focus from solely assessing the final product to evaluating the learning process—and the responsible, reflective use of AI within it:

- **Redefining academic integrity:** Develop clear guidelines for ethical use of AI in academic work. This means moving beyond a simplistic view of AI as a form of cheating, and instead framing it as a tool that can be used responsibly or irresponsibly. Transparency, attribution and critical engagement with AI outputs should be emphasized.
- **Process-oriented assessment:** Implementing assessment methods that capture the learning process, such as:
    - **Reflective journals/portfolios:** Students document their learning journey, including their interactions with AI, the prompts they used, the challenges they faced, and the insights they gained. This allows educators to see how students are using AI and to assess understanding of underlying concepts.
    - **Iterative assignments:** Projects that require students to show their work, including multiple drafts, revisions, and reflections on how they used AI tools and refined their outputs. This emphasizes the iterative nature of learning and problem-solving.
    - **Oral defenses/viva voce:** Students defend their work orally, demonstrating their understanding of the subject matter and their ability to think critically about the role of AI in their work. This allows for a more nuanced assessment of student learning and can help to identify instances where students may have relied too heavily on AI.
    - **Peer assessment:** Students evaluate each other's work, focusing on the quality of reasoning, critical engagement with AI, and ethical considerations. This promotes a sense of collective responsibility for learning and can help students develop their own critical evaluation skills.
- **Assessing prompting and teaching skills:** Directly assess students on their ability to effectively prompt and guide AI tools. This involves evaluating their understanding of how different prompts elicit different responses, their ability to refine prompts iteratively, and their capacity to critically evaluate AI-generated outputs. Students can be tasked with 'teaching' an AI to perform specific business-related tasks.
- **Promoting a growth mindset:** Encouraging students to view AI as a learning partner rather than a shortcut. Emphasizing that the goal is not to avoid using AI, but to learn how to use it effectively and ethically to enhance one's understanding and capabilities.

Ultimately, to borrow a line from Georgetown University professor, Jeanine W. Turner, the challenge for faculties is to reimagine learning and assessment in a context 'where the critique of the answer *is* the answer'[25]—and expertise is a property not of solitary humans, but of hybrid human-machine systems.

> **… the challenge for faculties is to reimagine learning and assessment in a context 'where the critique of the answer *is* the answer'**

25 Personal correspondence during the review process for this paper, reproduced with permission and thanks.

Practical recommendations

1. Integrate critical thinking elements into all courses

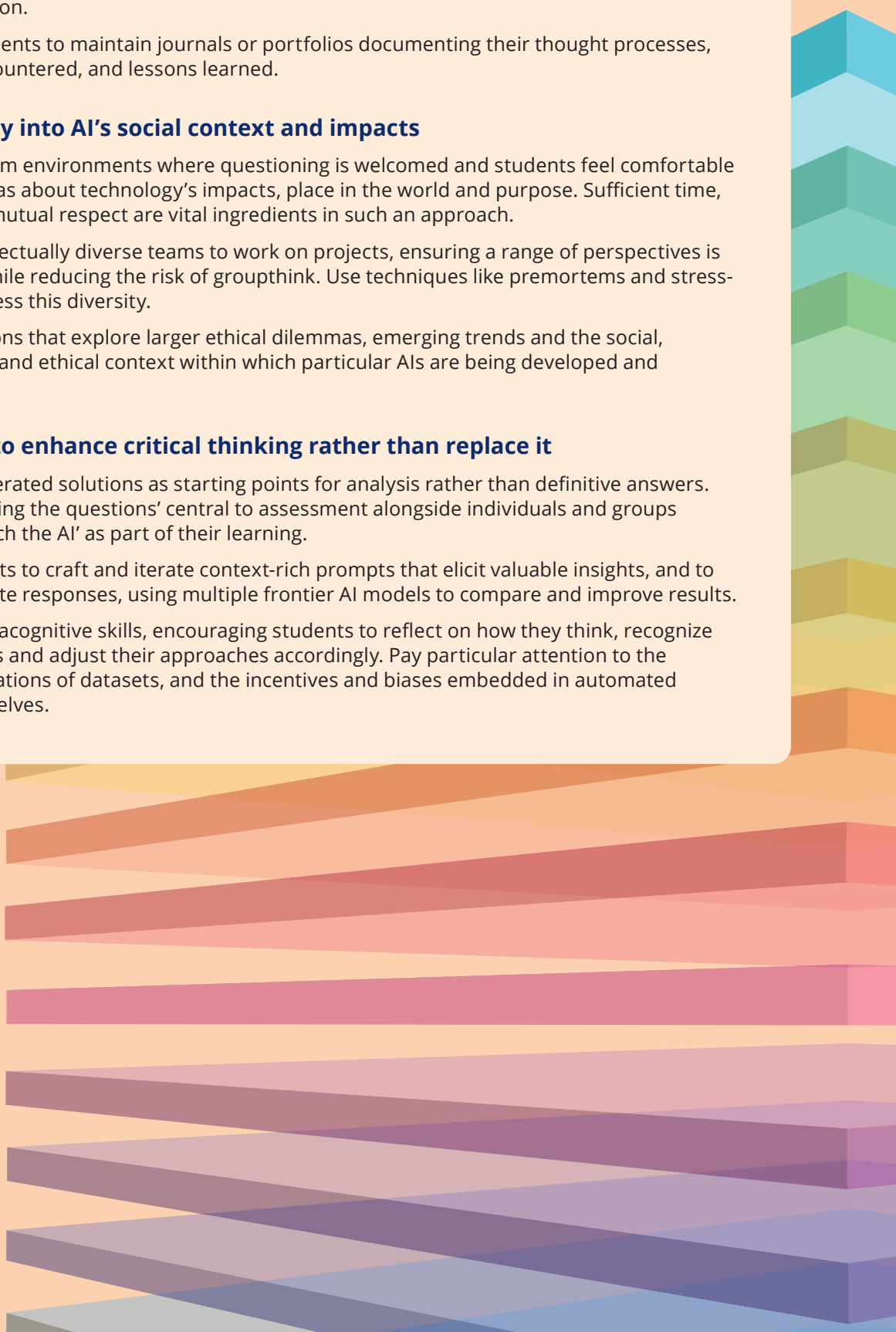
- Design assignments that require students to solve complex problems using both AI tools and critical analysis. Teach the fundamentals of reasoned arguments, well-evidenced explanations, and plausible theories.
- Combine theoretical instruction around human cognitive strengths and weakness with real-world scenarios that illustrate practical techniques for overcoming bias, groupthink and oversimplification.
- Encourage students to maintain journals or portfolios documenting their thought processes, challenges encountered, and lessons learned.

2. Foster inquiry into AI's social context and impacts

- Create classroom environments where questioning is welcomed and students feel comfortable challenging ideas about technology's impacts, place in the world and purpose. Sufficient time, attention and mutual respect are vital ingredients in such an approach.
- Assemble intellectually diverse teams to work on projects, ensuring a range of perspectives is represented while reducing the risk of groupthink. Use techniques like premortems and stress-testing to harness this diversity.
- Organize sessions that explore larger ethical dilemmas, emerging trends and the social, environmental and ethical context within which particular AIs are being developed and deployed.

3. Leverage AI to enhance critical thinking rather than replace it

- Present AI-generated solutions as starting points for analysis rather than definitive answers. Make 'questioning the questions' central to assessment alongside individuals and groups learning to 'teach the AI' as part of their learning.
- Instruct students to craft and iterate context-rich prompts that elicit valuable insights, and to critically evaluate responses, using multiple frontier AI models to compare and improve results.
- Emphasize metacognitive skills, encouraging students to reflect on how they think, recognize cognitive biases and adjust their approaches accordingly. Pay particular attention to the structural limitations of datasets, and the incentives and biases embedded in automated systems themselves.



It’s not enough to say we must keep humans in the loop. The question is *where, when and how* this should be done.

Part 3: Augment human abilities

Building on the centrality of critical thinking, we can envision AI as a 'cognitive catalyst' in our future relationship with technology. Much as a chemical catalyst accelerates reactions, AI can enhance our mental faculties, enabling us to process information, generate insights, and make better decisions. In general:

- People are sociable, empathetic, creative, critical thinkers who thrive on small amounts of actionable data, constructive dissent, rest, and variety.
- A heightened, sociologically literate focus on these 'soft' interpersonal skills is vital for defending and deepening the place of humans in the future workplace.

By contrast:

- ML systems excel at extrapolating from vast amounts of data, handled at immense speed and scale. They benefit from constant connectivity and predictability but struggle in circumstances with less or poorer quality data.
- They also find it challenging to self-correct, distinguish between accurate and merely plausible outputs, or reason rigorously and conceptually.

Ironically, it is humans rather than systems like LLMs that excel at logic and the meticulous use of evidence. While LLMs are readily able to generate multiple ideas, drafts, scenarios, and speculations, humans' metacognitive skills are essential for directing and appraising these.[26]

The architecture of involvement

As the concept of AI as cognitive catalyst emphasizes, the challenge isn't simply to monitor AI's outputs, but to design workflows and educational approaches that reflect the distinct capabilities of both humans and machines. It's not enough to say we must keep humans in the loop. The question is *where, when and how* this should be done. As author and AI risk expert, Cathy O'Neil argued in her November 2023 evidence on AI safety provided to the US Senate, the metaphor of a 'cockpit' can be valuable when thinking about human engagement with AI systems:

> *In considering the appropriate scope for monitoring/auditing AI systems and regulatory oversight, we often use the metaphor of a cockpit in an airplane. Planes are of course thoroughly tested before leaving the factory, and inspected before each flight. But we would never get on an airplane that had no cockpit. To fly safely, pilots need to monitor changing conditions, know when danger is imminent, and adjust accordingly. This is exactly what the cockpit enables.*[27]

The key is to design this relationship so that AI augments human agency, creativity, and judgment, rather than undermining or replacing them. A well-designed cockpit exemplifies this, acting as a cognitive catalyst for pilots. This in turn suggests three crucial principles for effective human-AI collaboration:

- **Operator training:** Pilots need to deeply understand both their own capabilities and the limitations of their instruments. Similarly, business leaders need to understand how AI systems work, their potential biases, and when to trust or question their outputs.
- **Interface design:** Just as a well-designed cockpit presents the right information at the right time, AI tools need to be designed to present insights in a way that is clear, actionable, and aligned with human cognitive processes. This involves careful consideration of what data to display, how to visualize it, and how to interact with it.
- **Balance of control:** Too much automation can lead to disengagement and deskilling, while too little can overwhelm users. The optimal balance allows AI to handle routine tasks and process vast amounts of data, while humans retain control over strategic decisions and handle ambiguous or exceptional situations. It's about finding the sweet spot where neither pilots nor business leaders are overwhelmed or asleep at the wheel.

26 For an interesting reflection of how this might apply to teaching practice, see the Oregon State University eCampus resource *Bloom's taxonomy revisited*. https://ecampus.oregonstate.edu/faculty/artificial-intelligence-tools/blooms-taxonomy-revisited-v2-2024.pdf and for some of the research informing it see, see Zaphir, L., Lodge, J. M., Lisec, J., McGrath, D. & Khosravi, H. (2024). *How critically can an AI think? A framework for evaluating the quality of thinking of generative artificial intelligence*. **https://arxiv.org/abs/2406.14769**

27 **www.schumer.senate.gov/imo/media/doc/Cathy%20O'Neil%20-%20Statement.pdf**

In the context of business education, the analogy is clear. We need to equip future leaders to interpret AI outputs, understand system limitations, and make informed decisions. And we need to pay close attention to the contexts that support rather than undermine such skills.

Data, ambiguity and complexity

The nature and quality of outputs from Machine Learning systems are dependent upon their training data. Working effectively with such systems therefore means not only understanding data but also knowing how to act on it wisely—and, in particular, understanding the kind of ambiguities and complexities that AIs themselves cannot. Embedding this kind of data literacy in business schools means integrating data-handling into as many aspects of learning and research as possible. Opportunities include:

- **Advanced analytics training:** Offer courses that teach statistical analysis, data visualization, and interpretation of AI outputs using tools like Python, R, or Tableau.
- **Real-world applications:** Engage students in projects requiring data-driven decision-making, emphasizing the integration of human judgment with AI assistance.
- **Simulated environments:** Use business simulations like Hubro that incorporate AI tools and assessment, allowing students to practice decision-making in settings mirroring real-world complexity.
- **Collaborative projects with industry:** Partner with organizations to provide students with real datasets and problems, fostering hands-on experience and industry relevance.
- **Ethical data handling:** Embed ethics into data practices, ensuring students understand the implications of data use and AI recommendations, particularly concerning privacy and bias.
- **Computational thinking frameworks:** Teach frameworks such as Wolfram's 'Define. Abstract, Compute. Interpret.' to help students conceptualize real-world problems in ways that make them tractable to solve with the support of computation.[28]

Example: A capstone project where students work with a retail company to analyze customer purchasing data using AI combined with a range of statistical tools, developing targeted marketing strategies while considering ethical implications.

All of the above must be done while heeding perhaps the most crucial lesson fields like aviation can teach: it's essential to analyze case studies of failure and error if you hope to avoid these in future. As *The Economist* bluntly put it in November 2024:

> *Productivity growth remains in the doldrums, far below where it was in the 1960s or 1970s. AI-boosters will tell you that the technology is transforming the fortunes of this company or that. Don't believe them.*[29]

Does this mean none of the hype is true? No. 2024's Nobel Prizes in physics and chemistry recognized breakthroughs where AI played a crucial role in analyzing complex data and accelerating discoveries.[30] Other striking examples of AI-driven innovation range from material science and medical imaging to biotechnology and healthcare logistics.[31] Across much of the broader economy, however, promised productivity gains have not yet materialized. This gap is not atypical so far as a general-purpose technology is concerned.[32] Indeed, it may be the case that measures such as 'productivity' neither capture current trends nor the most important areas of innovation.

28 See **www.computationalthinking.org/about/index.php.en** for details of the Wolfram approach

29 Williams, C. (2024, November 20). *There will be no immediate productivity boost from AI*. The Economist. **www.economist.com/the-world-ahead/2024/11/20/there-will-be-no-immediate-productivity-boost-from-ai** See also Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A. & Nicoletti, G. (2024). The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. OECD Artificial Intelligence Papers, No. 15. *OECD Publishing*. **https://doi.org/10.1787/8d900037-en**.

30 See **www.nobelprize.org/all-nobel-prizes-2024/**

31 For an in-depth picture of the gains, promises and disappointments of AI in the marketplace, see the State of AI Report produced by AI investor Nathan Benaich and Air Street Capital, **www.stateof.ai/**

32 See, for example, the analysis by ING in the report, AI productivity gains may be smaller than you're expecting' (2024, April 12). **https://think.ing.com/articles/macro-level-productivity-gains-ai-coming-artificial-intelligence-the-effect-smaller/**

Precisely because of this, however, business schools need to emphasize higher level questions about how AIs' impacts and potentials should be approached and assessed:

- How should qualitative or indirect benefits be accounted for?
- How can automation be compared and contrasted to augmentation?
- What mix of research methods might best capture changes within different domains?
- Within which sectors are significant trends most likely first to manifest?

Learning to think like a (social) scientist

Catalyzing cognition also means determining which analytical frameworks are most empowering. In an AI-rich environment, it is essential that leaders are comfortable operating amid uncertainty and complexity, making decisions with incomplete information and adapting to rapidly changing circumstances. This means adopting standards of proof, confidence, and evidence rooted in rigorous science, alongside methodologies adequate for navigating the intricacies of human behavior. In particular, learners and leaders need to know how to:

- Test rather than seek to confirm underlying assumptions
- Apply rigorous analytical methods and seek high quality evidence
- Honestly engage with real-world complexities and ambiguities
- Be prepared to change their minds and theories when the evidence changes

This emphasizes the fact that, given its sheer flexibility as a technology, a defining challenge for businesses (and, in turn, business schools) hoping to adopt Generative AI is rigorously demonstrating compelling use-cases. Here is technology expert Benedict Evans writing in September 2024 on this theme:

> *I keep thinking that the paradox of LLMs is that they present what looks like a 'human' interface, but that just makes them harder to use - what should you ask? As Steve Jobs said, it's not the customer's job to know what they want, nor to work out what to do with a new technology: it's the product's job to work that out...* [33]

The combination of data abundance and uncertainty inherent to an AI age cannot be navigated by declaring that present problems will be dissolved by ever-smarter systems. Rather, future leaders must be able to cast a discerning eye across specific evidence, claims and use-cases, balancing the use of quantitative data with qualitative insights and human intuition.

Ultimately, the challenge for tomorrow's leaders might be defined as determining, in an age where machines can answer almost any question instantly, which questions are worth answering in the first place. The question is not only what *can* be done, but what *should* be done—and how AI can catalyze the very best of the human.

Example: In an MBA market entry project, student teams use AI to analyze emerging market opportunities in terms of statistical trends, then critically evaluate its recommendations against local cultural and regulatory factors absent from training data but present in ethnographic and qualitative research. The resulting strategies demonstrate how human contextual understanding can transform AI analysis into practical insight—and the significance of a team's literacy across a variety of research methods.

... a defining challenge for businesses (and, in turn, business schools) hoping to adopt Generative AI is rigorously demonstrating compelling use-cases.

33 Evans, B. (2024, September 17). Benedict's *Newsletter*, No. 558, Free Edition. www.ben-evans.com/newsletter

Practical recommendations for business schools

Emphasize data-driven decision-making

- Develop case studies that position AI as a 'cognitive catalyst', requiring students to leverage AI's analytical capabilities while applying human judgment to interpret needs and opportunities—in the process exploring a future of AI/human hybrid teams and expertise.
- Design learning experiences where students analyze complex datasets, interpret AI-generated insights, and make strategic decisions using a rigorous approach to data—and actively questioning the uses and limitations of measures like productivity and efficiency.
- Create simulations where student teams use AI to analyze market trends, assess risks and make data-driven investment decisions, fostering evidence-based problem-solving.

Cultivate skills for navigating ambiguity and complexity

- Develop workshops and exercises around analyzing ambiguous or incomplete data. Encourage students to challenge assumptions, consider diverse perspectives, and synthesize human and AI insights.
- Train students to communicate effectively in uncertain situations, articulating the limitations of data and knowledge, and building consensus through dialogue and shared understanding.
- Have student teams analyze complex social issues with the help of AI, formulating recommendations while acknowledging uncertainty and drawing upon a range of research methods.

Prioritize strategic questioning and problem framing

- Emphasize the importance of applying scientific methodologies to complex business problems, formulating hypotheses that can be meaningfully tested—and developing rigorous criteria for failure and success.
- Teach students to rigorously define problem scopes, identify key variables, and explore the strengths and limits of datasets alongside the usefulness of different tools, interfaces, and metrics.
- Have student teams use AI to analyze complex business challenges, then actively reflect upon the strengths and limitations of both its insights and their own —and what making the best of both should mean.

Part 4: Lead through collaboration

Collaborative, interdisciplinary thinking is essential for addressing challenges that transcend traditional business boundaries. Indeed, AI systems themselves operate at the intersection of multiple disciplines, demanding a broad understanding of multiple domains if they are to be deployed effectively. Grappling with these complexities means acknowledging that the unintended consequences of AI and other complex systems define today's most urgent challenges—and that labelling them 'unintended side effects' is itself part of the problem. As John D. Sterman, director of the MIT System Dynamics Group, put it in a classic article:

> *There are no side effects—just effects. Those we expected or that prove beneficial we call the main effects and claim credit. Those that undercut our policies and cause harm we claim to be side effects, hoping to excuse the failure of our intervention. 'Side effects' are not a feature of reality but a sign that the boundaries of our mental models are too narrow, our time horizons too short.*[34]

To thrive in such an environment, future leaders need to:

- **Develop a systems thinking approach,** understanding how different fields and factors interconnect—and effects emerge from them in non-linear ways.
- **Encourage intellectual curiosity and openness** to ideas from diverse disciplines, with the capacity both to convene human experts and the capabilities of diverse automated systems.
- **Adapt quickly and actively to new technologies and methodologies** as they emerge, assessing their uses and risks in a critical, coherent manner rather than passively adopting system defaults without due consideration.

With AI promising narrow forms of expertise 'on tap' across fields ranging from coding and law to medicine and engineering, there's at least as much—if not more—value in people who can transcend traditional disciplinary boundaries and maintain cognitive flexibility.[35]

Reimagining collaboration for an AI age

To understand how to foster collaboration in an AI age, it's worth examining why and how humans collaborate in the first place. In the 2024 book, *The Social Brain,* evolutionary psychologist Robin Dunbar collaborated with leadership development experts Tracey Camilleri and Samantha Rockey to analyze the deep roots of human cooperation and cognition—and the fact that meaningful, mutually enriching social interactions are non-optional when it comes to high performance and thriving organizations:

> *All human organizations are social by nature, and our social propensities have very deep evolutionary roots… Humans are hardwired for connection and being part of a social group is critical.*[36]

Far from being 'soft' extras, the neurobiological systems that enable trust and cooperation took millions of years to evolve.[37] This has profound implications for how we structure teams and organizations in an age of global connectivity and AI-enabled scaling, not least when it comes to ensuring mutual trust, empathy, and respect—and the rich, reflective forms of communication that come when diverse perspectives are convened around a common purpose.

34 Sterman J. D. (2006). Learning from evidence in a complex world. *American Journal of Public Health*, *96*(3), 505–514. **https://doi.org/10.2105/AJPH.2005.066043**

35 See, for context, the increasingly impressive performance of frontier LLMs across measures of professional competence such as competitive programming, see **https://openai.com/index/learning-to-reason-with-llms/**

36 Camilleri, T., Rocket, S. & Dunbar, R. (2024). *The social brain: The psychology of successful groups*. Penguin (digital edition, locations 380 & 450).

37 Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*, 469–493. **https://doi.org/10.1016/0047-2484(92)90081-J**

As work like Dunbar's emphasizes, rather than treating collaboration as a distinct skill to be taught alongside others, it should be seen as the fundamental context within which learning and innovation takes place. Indeed, as management researchers Christine Moser and Nelson Phillips have highlighted in their research into the rise of 'symbolic machines,' *all* human cognition is enmeshed with both social and informational systems—and it makes little sense to assess these separately from one another:

> *In organizations, symbolic machines are everywhere, from systems that rate our performance, software that organizes our days, generative AI that writes texts and produces PowerPoint presentations, and recommendations and nudges that steer us toward particular choices and away from others. In short, symbolic machines change the very way that we live and work, both of which are inseparably linked with the symbolic.*[38]

In practical terms, acknowledging this entails business schools moving beyond a focus on traditional group projects and team-building towards what might be called 'collaborative intelligence environments'—spaces where human social cognition and AI capabilities can productively intersect. Key enablers include:

- **Psychological safety at scale:** Drawing on the work of researchers like Amy Edmondson in *Administrative Science Quarterly,* there is evidence that creating psychologically safe spaces for experimentation and learning becomes more challenging—and more crucial—as tools like AI introduce new forms of uncertainty and vulnerability.[39] Business schools must actively cultivate environments where students feel empowered to experiment openly with AI, failure is treated as a valuable source of learning rather than a mark of incompetence, diverse perspectives are actively sought out, and the boundaries between human and machine contributions are actively explored rather than assumed.
- **Cross-pollination of ideas:** AI's ability to process vast amounts of information across domains creates new opportunities for interdisciplinary insight. However, meaningful synthesis still requires human judgment and creativity. Business schools should align both their teaching and operations with structural forms of interdisciplinary collaboration, fostering the development of 'translation' skills—the ability to communicate complex ideas across different domains and stakeholder groups.
- **Ethical frameworks for collective decision-making:** As AI systems become more powerful, the ethical dimensions of decision-making become more complex. Collaborative approaches to ethical reasoning become essential. This entails developing shared vocabularies and frameworks for discussing ethical implications, addressing potential biases and unintended consequences, and building capacity for collective moral reasoning and judgment.

... business schools moving beyond a focus on traditional group projects and team-building towards what might be called 'collaborative intelligence environments'

38 Phillips, N. & Moser, C. (2024). The biological basis of the symbolic: Exploring the implications of the co□evolution of language, cognition and sociality for management studies. *Journal of Management Studies*, *61*(8), 3793–3823. **https://doi.org/10.1111/joms.13037**

39 Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. **https://doi.org/10.2307/2666999**

Transformative leadership through integration

The challenge of leadership in an AI age extends far beyond technical competence or even collaborative facilitation. As examples of innovative business schools like Stanford GSB and INSEAD demonstrate, true transformation occurs when leaders can integrate diverse perspectives, bridge disciplinary boundaries, and create environments where both hard and soft skills are valued and developed. Indeed, many established models of 'leadership'—whether transformative or otherwise—are likely to be tested to the point of fracture by present challenges.[40] Consider the paradox that many students enter programs hoping to emulate figures like Sheryl Sandberg or Elon Musk. Yet the most successful alumni may tread entirely unexpected paths—discovering new passions and possibilities along the margins between disciplines. Today's leaders must be able to:

- Design and maintain collaborative systems that span human and AI components
- Navigate the complex interplay between formal structures and informal social networks
- Cultivate innovation while maintaining ethical guardrails and operational effectiveness
- Balance automation with the need for meaningful human connection and development

Business schools can develop these skills by embracing interdisciplinary approaches and, in particular, increasing the permeability of the boundaries between their own practices and those of the wider world:

- Integrating cross-disciplinary coursework that combines business with the arts, humanities, STEM, and social sciences.
- Creating flexible aspects of curricula that allow students to explore (and share) diverse interests and tailor their learning experience—and faculty to co-teach such modules.
- Encouraging collaborative projects between students from different academic backgrounds, with an emphasis on the skills that facilitate such collaborations.
- Building participation with industry partners into such projects at appropriate stages, alongside establishing innovation spaces for co-design of curricular elements with students, faculty and industry.
- Anticipating potential challenges amongst faculty by giving them ownership over interdisciplinary content development, maintaining lines of communication, and providing support and resources for expanding their own interdisciplinary knowledge.

What the authors H. James Wilson and Paul R. Daugherty in a recent issue of *Harvard Business Review* call 'fusion skills'[41] become vital for using technology well: skills that help people pool the insights of diverse experts, systems, and data sources. This is a team sport. The era of the leader who knows or controls everything is over. Instead, the key lies in understanding the capacities and competences of hybrid human-machine systems—and guiding their application with imagination, rigor, and moral clarity.

As examples of innovative business schools like Stanford GSB and INSEAD demonstrate, true transformation occurs when leaders can integrate diverse perspectives, bridge disciplinary boundaries, and create environments where both hard and soft skills are valued and developed.

40 For example, the BANI framework (Brittle, Anxious, Non-linear, Incomprehensible), first introduced by futurist Jamais Cascio in his 2020 article Facing the Age of Chaos, offers a contemporary lens on today's increasingly complex business environment. It is proposed as an updating of the long-standing VUCA model (Volatility, Uncertainty, Complexity, Ambiguity), developed in the late 1980s by the US Army War College to describe the post-Cold War world. **https://medium.com/**@cascio/facing-the-age-of-chaos-b00687b1f51d

41 Wilson, H. J. & Daugherty, P. R. (2024, September–October). Embracing gen AI at work. *Harvard Business Review.* **https://hbr.org/2024/09/embracing-gen-ai-at-work**

Practical recommendations

1. Create integrated learning environments

- Set up physical innovation labs where interdisciplinary teams can work on real-world scenarios and industry problems, with dedicated spaces for both AI-assisted analysis and 'screen free' human-to-human dialogue.
- Establish regular forums where students share and critique their experiences of integrating AI tools with human collaboration, focusing on both successes and instructive failures.
- Create connections between student groups working on similar challenges across different business schools, building networks of shared learning about effective collaboration in an AI age.

2. Foster deep integration across disciplines

- Create integrated learning modules that combine technical AI literacy with humanities and social science perspectives, helping students develop what Wilson and Daugherty term 'fusion skills'—the capacity to combine human and machine insights while maintaining ethical awareness and critical distance.
- Establish ongoing partnerships between technical and non-technical university departments to co-develop curriculum that addresses both the *how* and *why* of AI deployment, grounded in real cases of both success and instructive failure.
- Design collaborative projects that require students to navigate between quantitative AI analysis and qualitative human factors, building the kind of cognitive flexibility crucial for future teams and leaders.

3. Cultivate leaders able to embrace uncertainty and complexity

- Train students in practical methods for building psychological safety in teams working with AI, including how to encourage experimentation and learn productively from failure.
- Develop frameworks for responsible AI implementation that emphasize Dunbar's insights about human social cognition and trust alongside technical capabilities.
- Foster leadership approaches that embrace uncertainty and complexity through structured experimentation, helping students move beyond simplistic 'AI adoption' to thoughtful integration of human and machine capabilities.

Implementing the DUAL framework

Implementing the DUAL framework is a long-term process that will require ongoing effort, adaptation, and collaboration. Recognizing that business schools have different starting points, resources and constraints, this paper proposes a tiered approach to implementation:

Tier 1: Short-term/low-resource adaptations

These are changes that can be implemented relatively quickly and easily within existing structures and with minimal resources:

- **Integrating AI literacy into existing courses:** Adding modules or discussions on AI capabilities, limitations, and ethical implications to existing courses across the curriculum.
- **Using AI tools for specific tasks:** Incorporating AI tools into assignments in a focused way, such as using a Large Language Model for brainstorming or an AI-powered research tool for gathering information.
- **Developing clear guidelines on AI usage:** Creating clear and transparent guidelines for students on the ethical use of AI in coursework, emphasizing transparency, attribution, and critical engagement.
- **Faculty development workshops:** Offering short workshops or training sessions for faculty on how to integrate AI into their teaching and assessment practices.
- **Promoting discussion and dialogue:** Creating forums for faculty and students to discuss the implications of AI for business research and education, to share best practices, and to debate AI's wider social, ethical and environmental impacts.

Tier 2: Mid-term/moderate-resource changes

These changes require some structural adjustments and a moderate level of investment:

- **Developing new modules or courses:** Creating new modules or courses focused specifically on AI ethics, AI strategy, or data analytics with AI.
- **Creating interdisciplinary projects:** Designing projects that require students to work in interdisciplinary teams, combining business knowledge with technical skills to address AI-related challenges.
- **Revising assessment methods:** Implementing process-oriented assessment methods, such as reflective portfolios, iterative assignments, and oral defenses, to better capture student learning in an AI-enabled environment.
- **Building partnerships with industry:** Collaborating with industry partners to provide students with real-world data sets and AI-related business problems to work on.
- **Investing in AI tools and infrastructure:** Providing students and faculty with access to relevant AI tools and platforms, and ensuring that the necessary IT infrastructure is in place.

Tier 3: Long-term/high-resource transformations

These are more fundamental changes that require significant institutional commitment and investment:

- **Redesigning curricula:** Fundamentally rethinking the curriculum to integrate AI across all aspects of business education, emphasizing the development of critical thinking, collaborative and ethical skills, alongside technical proficiency.
- **Creating collaborative learning spaces:** Designing physical and virtual spaces that facilitate interaction between students, faculty, and AI tools, fostering a culture of experimentation and co-creation.
- **Developing new performance metrics:** Revising faculty and institutional performance metrics to value teaching innovation, interdisciplinary collaboration, and the development of human-centered AI skills.
- **Fostering a culture of continuous learning:** Creating a culture that embraces lifelong learning and adaptation, recognizing that the skills needed to thrive in an AI-driven world will continue to evolve.
- **Promoting research on AI in business education:** Supporting research on the pedagogical implications of AI, the ethical and societal impact of AI in business, and the development of new assessment methods for an AI-enabled world.

Conclusion

The most significant questions posed by AI to business schools are not technological. They are about what it means to learn, lead and create value in an age where novel, evolving forms of machine 'intelligence' are abundant. Similarly, addressing these questions demands expertise in those areas machines find hardest to navigate: weighing up fundamental values and assumptions; determining which problems are worth solving; articulating the intricacies and ambiguities of human experiences, needs, and relationships. In an age of artificial intelligence, success more than ever depends upon distinctively human capacities.

For business schools, this suggests a future where value lies not in delivering answers or developing narrow expertise, but in teaching students to cultivate judgment, empathy and self-reflection—then to test and translate their insights within machines' data-driven realm. Implementation will demand patience, experimentation, and honest engagement with constraints. But the alternative poses greater risks. The goal is not to race against machines, but to investigate how human and artificial intelligence can complement one another in the service of meaningful ends—and what it means for individuals, teams and organizations to thrive in information environments increasingly interwoven with human minds.

Perhaps most importantly, it means remembering that technology's ultimate purpose is not to replace human judgment, but to create contexts within which it can flourish.

“The most significant questions posed by AI to business schools are not technological. They are about what it means to learn, lead and create value in an age where novel, evolving forms of machine 'intelligence' are abundant.

About the author

Dr Tom Chatfield is a British tech philosopher, author, advisor, and speaker. He writes, broadcasts and consults internationally, with a special interest in critical thinking, AI, and future skills. His recent work includes designing and presenting new critical thinking and AI business courses for *The Economist;* running international seminars on the critically discerning use of AI; and co-creating a prototype AI cognitive co-pilot.

An experienced Chair and NED, he's currently a member of the British Library Advisory Council, and Associate at the interdisciplinary think-tank, Perspectiva. His books exploring digital culture—including *Wise Animals, How To Thrive in the Digital Age, and Fun Inc.*—are published in over thirty languages, while his critical thinking textbooks, published by Sage, are used by universities, institutions, and companies around the world.

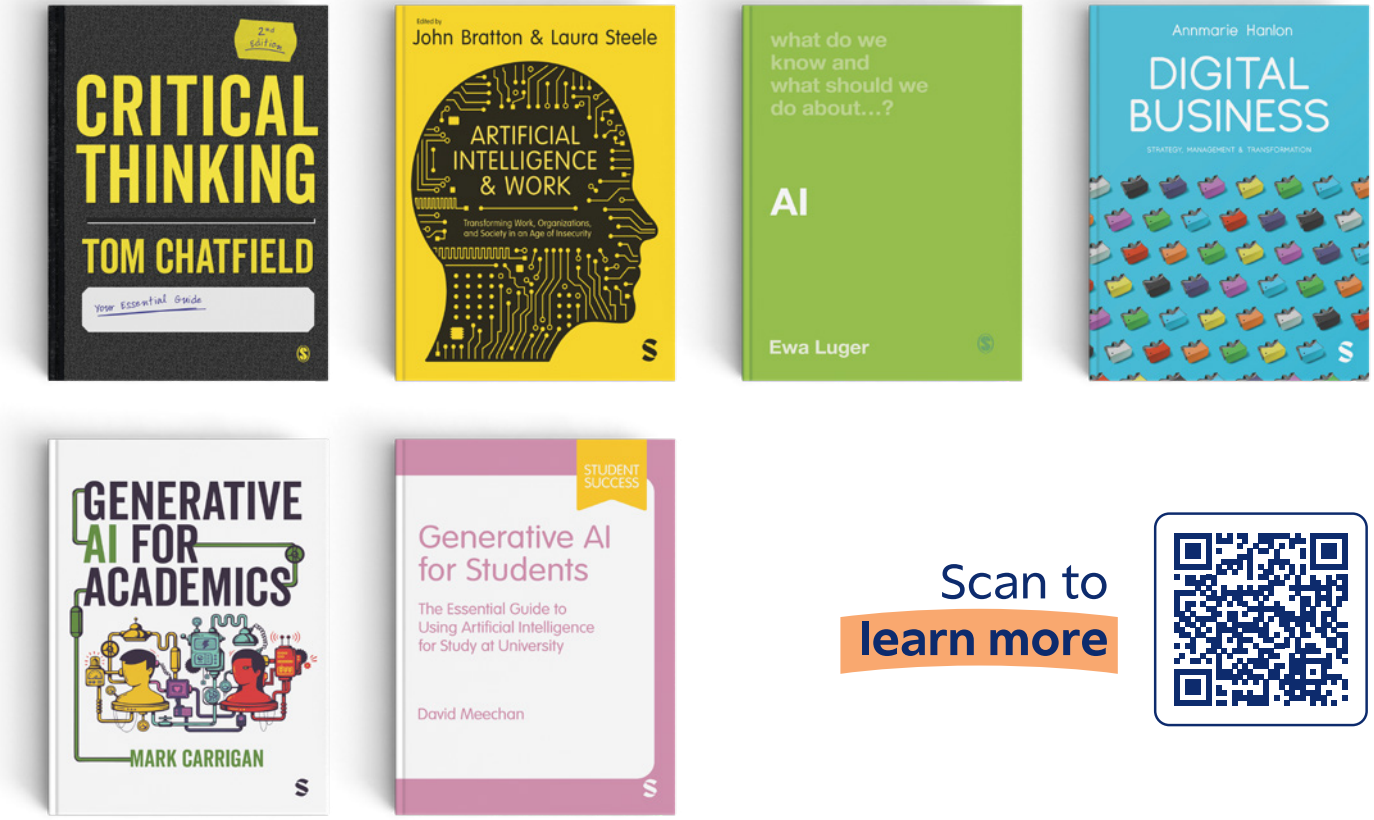
Over the past decade, Tom has developed award-winning online courses for universities and businesses, and been a guest faculty member at the Said Business School and visiting associate at the Oxford Internet Institute, both at the University of Oxford. He has collaborated with organizations including Allianz, the BBC, Google, and the UN. Past speaking appearances include TED Global and the US National Academy of Sciences. A launch columnist for BBC Future, he writes widely in the international media and guest lectures at universities in the US, UK and Europe.

Acknowledgements

I am indebted to this paper's advance readers and reviewers: Tracey Camilleri, Dr Timo Hannay, Professor Dr Jasmin Mahadevan, Professor Neil Maiden, Professor Christopher Mairs, Dr Christine Moser, James Mustich, and Professor Jeanine Turner. Their insights and generous feedback have improved it immeasurably, as has the guiding expertise of Matthew Waters at Sage, to whom I also owe the suggestion of the DUAL acronym. Thank you.

In line with my own recommendations, I have 'talked through' a wide range of points included here with the latest 'frontier' versions (as of January 2025) of Gemini, Claude, DeepSeek and ChatGPT, as well as challenging them to pick holes in my ideas, highlight where I'm being vague or verbose, and play the part of potential readers. The quality of the responses I've received has been mixed. Often, I find the kind of prose that LLMs propose to be turgid and strewn with jargon, and their assessments of my own writing to be unreliable—or unsympathetic to my philosophical inclinations. At the same time, they can be wonderful explainers of ideas, co-readers of resources, and spotters of non-sequiturs. I would like to think that my access to LLMs has made this paper better. Certainly, it has made it easier to hold and structure diverse thoughts, check through the logic of particular points, and tease out weaknesses. However, it's worth emphasizing that no LLM's analysis has come close to those provided by the expert human readers above, thanks to the machines' lack of an individual perspective, real-world experience, contextual awareness, and deep knowledge of relevant literature. It may be the case that some readers will prefer a briefer, simpler version of this paper as digested by an LLM. I sympathize with the desire for brevity. But I hope that, for those prepared to work their way through it more closely, there is value in its details and diversions—and enough of a personal perspective to be worth disagreeing with.

Explore AI and related skills in Sage books & journals



Scan to **learn more**

Scan to **learn more**

Explore AI and related skills in Sage learning resources

Sage Business

Sage Business Data Decisions bridges the gap between business fundamentals and practical applications, helping students build the competencies needed to excel in a competitive job market. It enables them to evaluate and apply data across various business scenarios using critical thinking, and enhances their ability to communicate findings.

Sage Skills helps students as they transition into university study and as they prepare for post-graduation in the professional world. Supporting the development of a wide range of skills, the suite builds confidence while fostering critical thinking in real-life situations.

Sage Campus

Free course

The Art of ChatPGT Interactions is a free course designed to help you build the skills and confidence you need to get the most out of generative AI tools such as ChatGPT. Under the guidance of AI expert Leo Lo, you'll master the art of prompt engineering, deepen your understanding of how ChatGPT works, and learn how a systematic approach can help you get better outcomes.

Scan to **learn more**

Hubro Simulations

Our award-winning simulations allow students to apply real-world business skills in a risk-free environment, whether individually or in teams, in the classroom or online. Save time on assessments and provide students with personal feedback with our new AI Grading tool.

For further information about any of the above, email **business.management@sagepub.com**

www.sagepublications.com

2025