Artificial Intelligence and Digital Learning: Architecture, Hallucinations, Information, Findability, The First Rung, and The Arts of Inquiry

Gardner Campbell
Virgina Commonwealth University

Charles Dziuban
University of Central Florida

Colm Howlin
Independent Data Scientist

Mark Smith
TGS Energy Data and Intelligence

Authors in alphabetical order

Abstract
The authors address Artificial Intelligence (AI) powered by Large Language Models (LLM) and its relationship to learning in contemporary education. Initially they explain the underlying functionality of AI using transformer architecture, embedding, and tokenization to create language symbolism. Next, they discuss the transformed search concept and how scale-free networks and power law distributions portray information sources dominated by AI hubs that couple and decouple digital learning resources. They contend that Artificial Intelligence will replace bottom- and entry-level jobs by removing a foundational rung of new graduates' career development. This shift, termed the *answer machine*, will impact graduates, industry, and education, creating an urgent need to mitigate the risks and leverage the opportunities AI presents. Finally, they consider potential consequences to human creativity, insight, wisdom-related knowledge, and the arts of inquiry in an AI-driven world where we face the danger of losing the ability to think about thinking. Artificial Intelligence can lead to a remarkably improved culture of education. At this time, however, with our seriously limited understanding of its long-term implications, AI presents more questions than answers, so the immediate need is to suggest a context for the best-informed and most important questions to emerge and receive robust consideration at all levels of society.

Keywords: Artificial Intelligence; digital learning; information structure; entry level jobs; inquiry process

Campbell, G., Dziuban, C., Howlin, C., & Smith, M. (2026). Artificial Intelligence and digital learning: Architecture, hallucinations, information, findability, the first rung, and the arts of inquiry. *Online Learning, 30*(2), 1-30. https://doi.org/10.24059/olj.v30i2.6050

Introduction: How Artificial Intelligence Works

As Generative AI powered with Large Language Models (LLMs) impacts society, it becomes essential to understand the influence of architecture, training, and alignment processes. They can exert a profound and often undetermined influence on knowledge dissemination, education, and learning (AAC&U, 2025; AAUP, 2025; Dubois, 2024; Robert Ovetz, 2025). Therefore, understanding how these systems are constructed and address "knowledge and values" provides insight into their capabilities and issues. While avoiding excessive technical depth, this section provides an explanation about the core components and training procedures that underpin modern LLMs.

Transformers, Attention, and Embeddings

Most modern LLMs are built upon transformer architecture, a deep learning approach introduced in 2017 (Vaswani et al., 2017). The original development proposed an encoder-decoder structure. Here, we will concentrate on the decoder component because it is the element that powers the LLMs behind applications like ChatGPT and has fueled the generative AI revolution.

At its foundation, the model takes input text and creates an initial embedding through a process known as "tokenization" (Pilehvar and Camacho-Collados, 2021). This takes words and subwords and converts them into numbers. Short common words like "the" are usually single tokens, while longer or more complex words get broken into multiple pieces. For example, the word "understanding" might be split into "under" and "standing."

After tokenization, each word gets converted into a vector or embedding. These are numeric representations of the token. Words and tokens that have similar meanings will have like numeric values, and words that are very different from each other will have different numeric positions. For example, we might expect "king" and "queen" to have similar values and "king" and "calculus" to be considerably different.

The models interact with these numeric tokens rather than the input text. The LLM processes them, much like we would read the words in the input text. As it processes (or reads) each token, it builds up a representation of what the text means. There are various mechanisms it uses for this, "self-attention" being the most prominent. Not all tokens in the sequence are equally important for understanding the current one under consideration and attention allows the model to understand which *previous* input elements are the most relevant to the *current* one and where it should focus. This self-attention mechanism is an important breakthrough in the development of LLMs, with the original paper titled "Attention is All You Need" (Vaswani et al., 2017).

Once it has read all the input tokens and established its internal representations, the final step assigns a probability to every possible next *token*. From here, it can select most likely candidates and assigns the next token in the sequence. There are several ways of selecting tokens from the candidates, with the highest probability being the simplest. Most models allow the introduction of some variation using a parameter known as "temperature." This term is borrowed from the simulated annealing method that allows users to introduce some randomness into the process. Many think of this as controlling the model's freedom or "creativity." The higher the temperature, the freer the model is to select from the most probable next candidates.

With the token selected, it is appended to the input, and the process is repeated to select the next one in the sequence. The model stops once it generates a special "EOS" or end-of-sentence token. Fundamentally, a Generative Model repeatedly predicts and appends tokens in a sequence.

Pre-training and Fine-tuning

The above section outlines what the process does, but input data trains the model to do this. These mechanisms have thousands, millions, and increasingly billions of associated parameters. Each one controls some aspect of how the model functions. With so many parameters, it is impossible to understand or specify them. The only way they can be set is for the model to "learn" them from data in the training process.

The sequence begins by setting the parameters at random. Then, by supplying inputs and comparing them to expected outputs, the parameters are adjusted by the training process so that the predicted output is brought closer to the expected result. This is repeated over *trillions* of tokens to train the largest models. For modern prompt-driven LLMs, the process is typically broken down into a two-stage training paradigm: pre-training and fine-tuning.

Pre-training is the initial, computationally intensive phase where the model learns general language understanding, grammar, common sense knowledge, and world facts. This is achieved through "self-supervised learning" (predicting the next or missing token/word) on vast, diverse datasets, often comprising trillions of tokens scraped from the internet (Guo et al., 2024). The output is called a pre-trained model. It can understand the inputs in the languages it was trained in, and it can predict the next token, but it cannot reliably accomplish more complicated tasks that we see modern LLMs achieve. That is where fine-tuning enhances the process.

Fine-tuning involves the pre-trained model being adapted or specialized for specific downstream tasks. Typically, this involves further training on smaller, more curated, and often labelled datasets relevant to the target application. Generally, these tasks are more complex (such as writing code or summarizing a document) (Xia et al., 2024).

Recently, a special class of fine-tuning has emerged called Reinforcement Learning from Human Feedback (RLHF). In this variation, the dataset used to fine-tune the model is based on human preferences (Zhang et al., 2024). This has been shown to increase the quality of outputs dramatically. However, a key point is that RLHF aligns the model to produce outputs humans prefer, not necessarily with factually correct answers. The system learns to replicate judgments made by these individuals (Liu, 2023), potentially introducing hidden biases.

Reliability: Errors and Hallucinations

A significant challenge for the use of LLMs in all applications, and in particular education, is "hallucinations." This is the LLMs' tendency to generate inaccurate or entirely fabricated information (Meade and Jano, 2024). This issue is attributable to fundamental design and training and poses considerable risks in settings where factual accuracy and trustworthiness are paramount.

Hallucinations can sometimes be obvious and caught by users; however, in contexts where the user has insufficient knowledge about the subject or task at hand, identifying these mistakes can be difficult, if not impossible. Often, this is the case with students' use of Gen AI. Coupled with the confidence with which these models speak, the results appear plausible and coherent but can be factually incorrect, nonsensical, irrelevant, or entirely fabricated (AAUP, 2025; Guo et al., 2024).

Several factors contribute to the occurrence of hallucinations:

1. The vast datasets used for pre-training inevitably contain errors, biases, outdated information, and inconsistencies.
2. Incomplete data on certain topics can also lead models to "fill in the gaps" with plausible but incorrect information (Guo et al., 2024), due to the RLHF fine-tuning goal to generate an output humans like as opposed to one that is correct.
3. This probabilistic generation process inherently allows for deviations from factuality (Guo et al., 2024).
4. A process known as overfitting (putting too much emphasis on specific patterns in the training data) can also contribute (Guo et al., 2024).
5. Most LLMs lack direct access to real-time information or a continuously updated external knowledge base (Zhang and Zhang, 2025). Its knowledge is essentially frozen at the time of training.
6. Ambiguous or poorly phrased prompts can lead the model to misinterpret the user's intent and generate irrelevant or incorrect responses (Guo et al., 2024).

Despite the ongoing efforts of developers to address these issues, hallucinations persist, leading some researchers to contend that they might be an inevitable structural feature of current LLM architectures (Banerjee et al., 2024).

Rapidly Developing Evolvement

Advancements on the horizon could significantly impact and strengthen the impact of AI. The trend expands past purely text-based towards multimodal models, capable of processing and generating text, images, audio, and video (Wadekar et al., 2024), with the possibility of personalizing content for learners.

Another key development likely to impact all Gen AI use is the work to improve model reasoning. This is sometimes referred to as "emergent" in LLMs, since it wasn't trained explicitly into the model but emerged due to the training data and architecture (Berti et al., n.d.). Developers are working to improve this capability with future iterations expected to exhibit more sophisticated logical thinking, enabling them to solve complex problems, understand cause-and-effect relationships, and make more informed decisions. This includes advancements in areas like multi-step reasoning, where the AI can break down a problem into smaller parts (Wang et al., 2024), and counterfactual reasoning, where the AI can understand "what if" scenarios (Alihan et al., 2025).

Another significant trend in the development of reasoning is the rise of Agentic AI, where systems move beyond passive responses to take proactive actions and make autonomous decisions (Purdy, 2024). The models use their reasoning capabilities to break down problems and arrive at results. The agent can function with a human in the loop or completely autonomously. Many industries have these autonomous agents in production, with several well-known examples such as Deep Research from OpenAI (Openai.com, 2025), and Manus (Manus.im, 2025). At the moment, the future of Artificial Intelligence appears unbounded, and progressing at a pace rarely encountered in the history of evolving technology.

Ambient Findability: Information Wayfinding in the Age of Artificial Intelligence

Twenty years ago, Peter Morville published *Ambient Findability* (2005), a definitive resource for information search and curation. He explained that findability coalesces navigation, labelling, and information architecture into a comprehensive knowledge acquisition model:

1. The ability to navigate and find information in a saturated environment.
2. The necessity of information literacy to evaluate unfiltered data.
3. Understanding the disproportionate distribution of information.
4. Understanding that information constitutes a complex network that cannot be readily disaggregated.

Additionally, he integrated the notion of “wayfinding” where search moves from one source to another in an iterative progression.

Wayfinding is a fancy word for a series of things people know and do to get from one place to another, inside or outside. Wayfinding can be a snap or an onerous take, depending on the person, the environment, and the situation. You can think of wayfinding as a five-step process. It starts with knowing where you are. It means knowing your destination, following the route to your destination, being able to recognize your destination and finding your way back to your starting point (Morville, 2005, p. 17).

Wayfinding is an accurate description of our contemporary educational environment. Teachers orient their way through their courses no matter how organized they are at the onset. Students orient their way through those same courses and together education becomes more than the sum of its individual parts like the emergence and self-organization of complex systems. However, modern information communication technologies, by interacting with each other, reframe Morville’s definition of wayfinding by limiting human involvement (Floridi, 2014). This relates to the Bates (1989) berry-picking search that is a non-linear search paradigm and corresponds to Wikipedia—the democratic information source that features open editing and continuous updates. According to Morville, the core principle of ambient findability is its potential for making information available anywhere at any time.

Although the book was an authoritative resource at the time, one of its significant contributions was its resonance with literature such as the chestnut tree in *Jane Eyre* (Brontë, 1847), the green light in *The Great Gatsby* (Fitzgerald, 1925) and the train in *Anna Karenina*

(Tolstoy, 1878), foreshadowing what was to come in the world of ubiquitous connectivity. However, the notion of foreshadowing becomes difficult in the context of artificial intelligence that is developing at an exponential rate, although that growth will moderate much like population growth modeled by the logistic map.

Morville did acknowledge the growing search potential of artificial intelligence (AI). However, today we find ourselves working with systems that alter how we engage with knowledge. We no longer actively search because information is served up as we ask by a metaphorical ATM. In the hyperconnected AI environment, the boundary between machine and human intelligence is blurring and changing what it means to know something (or even how). Our interaction becomes less mindful because we cede considerable human control to automated mediation. As a result, we live with a knowledge environment that must balance accessibility with intellectual agency.

Scale Free Networks and Asymmetric Power Law Distributions: Why They Matter

Contemporary science and technology influence every aspect of our lives. One effect is the formation of networks: for example, the internet, mobile phones, and social media (Harari, 2024). The most influential, however, are scale-free networks that develop as self-organizing entities in the absence of a central controlling mechanism (Caldarelli, 2007; Mitchell 2009). They feature a small number of hub nodes with many connections while most other nodes have far fewer pathways. Airline networks exhibit scale-free characteristics, as do social media sites where influencers dominate the connections. One explanation is the preferential attachment phenomenon where nodes entering the network connect to already established hubs (Barabási, 2016). Scale-free networks are remarkably resilient and information spreads rapidly throughout them. Additionally, they provide insight into the functioning of complex systems like the internet. However, the most important characteristic of these structures is that they conform to a power law distribution, sometimes called the long tail, where there is a piling up of occurrences on the left side with a steep decline going to the right (Newman, 2005; Anderson, 2006; Dziuban et al., 2023). This represents the marked asymmetry in the distribution of information (Taleb, 2007; Taleb, 2012; Taleb, 2018). There are multiple examples:

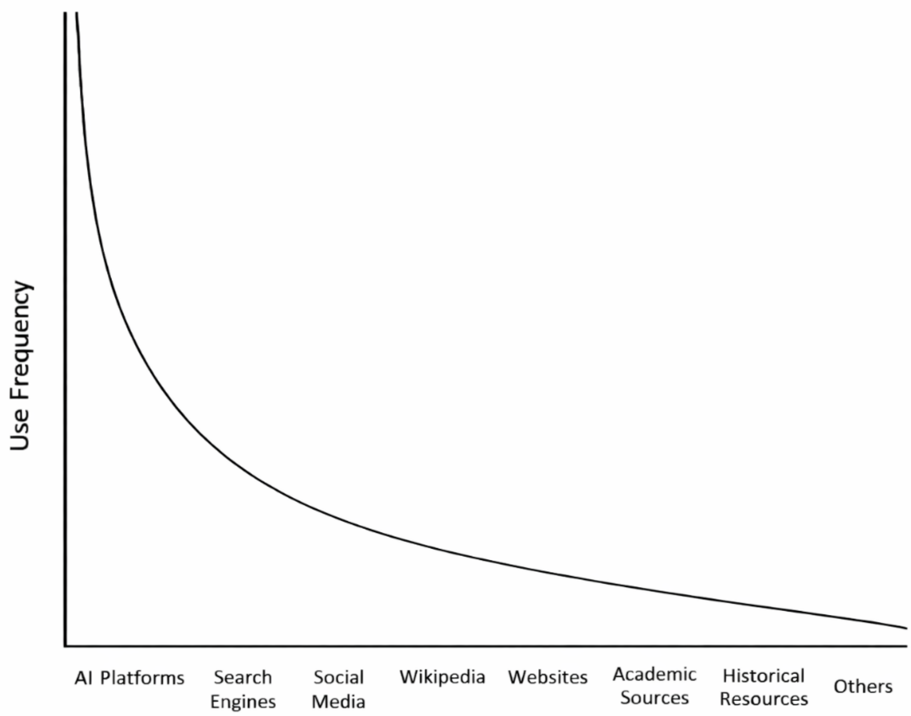
1. city populations
2. incomes
3. the size of corporations
4. word frequency in languages (Zipf, 1949)
5. internet traffic
6. academic article citations

In science, network thinking provides a novel language for expressing communalities across complex systems in nature, thus allowing for insights from one area to influence other, disparate areas. In a self-referential way, network science itself plays the role of hub—the common connection among otherwise far-flung scientific disciplines (Mitchell, 2009).

The question becomes then how do scale free networks and power law distributions relate to teaching, learning, and information search in the world of ambient findability, wayfinding, and artificial intelligence. If one examines the use of contemporary information sources, their distribution conforms to the scale-free power law. Consider Figure 1.

Figure 1

A Thought Experiment for the Power Law Distribution of Information Sources



If this thought experiment is correct, it represents a shift in the information process—that is, the information seeker becoming more passively receptive. The emergence of large language models as the dominating information source represents what Gladwell termed a tipping point (Gladwell, 2024), Kuhn labeled a paradigm shift (Kuhn, 1962), and Page identified as a phase transition (Page, n.d.). Examining figure one, we see that the flatter part of the curve represents a more slowly evolving search process. However, as the curve begins to accelerate, we see a rapid transition from Wikipedia to AI. The power law information source curve has two other noteworthy characteristics. Firstly, the sources listed on the horizonal axis are general prototype categories that denote large numbers of subcategories. Secondly, the graph is a static representation of dynamic processes where the categories change order or appearance, but on the left side of the horizontal axis the dominance of AI is unlikely to change soon. Figure 1 changes the adage "knowledge is power" to "the ability to use knowledge is power"—a dramatic transformation in the educational landscape. However, the modal portrayed in figure one is not yet validated and therefore should be considered a hypothetical construct that requires long term rigorous assessment and verification. That said, there are examples of this delayed validation in science: general relativity, the Higgs field, natural selection, continental drift, the cosmic microwave background, and exoplanets. Delayed validation is also the case in education and learning theory: for example, constructivism, critical period for language learning, mastery learning, and spaced repetition.

Artificial Intelligence and a Restructuring of the Information Framework

> The seed of the current revolution is the computer. Everything else from the internet to AI is a byproduct. (Harari, 2024, p. 193)
> Future generations will never know what an exclusively analogue, pre-digital reality was like. We are the last generation to have experienced it. (Floridi, 2019, p. XI)

These quotes from two leading thinkers in artificial intelligence either foreshadow the future or document the history of the information age. As Harari (2024) describes historical perspective, the narrative of history often becomes more important than the actual event. For instance, Rosa Parks and the Montgomery bus boycott, Paul Revere's ride, or Rosie the Riveter. These events were important in and of themselves but pale in comparison to the narratives that developed in the years following them. Acknowledging the long developmental history of large language models, in a remarkably brief time, the narrative has far surpassed a single event like the release of ChatGPT. The story of artificial intelligence has developed so rapidly that something remarkable has happened. Eric Hoffer, an American dock worker and moral and social philosopher, wrote about the nature of what he termed mass movements and the psychology of individuals within them. In his book "The True Believer: Thoughts on the Nature of Mass Movements" (1951), Hoffer describes "believers," a type we now find among those who are deeply involved in a movement like large language models. The pervasive media reach heightened by global interconnectedness accelerates the worldwide conversation, enthusiasm, skepticisms, and curiosity about artificial intelligence. The narrative is ubiquitous, compelling, and dominating.

Coupling and Decoupling

Floridi (2019) describes the process thus:

> The digital "cuts and pastes" our realities both ontologically and epistemologically. By this I mean that it couples, decouples, or recouples features of the world (ontology) and therefore our corresponding assumptions about them (epistemology) that we thought were unchangeable. It splits apart and fuses the 'atoms' of our 'modern' experience and culture so to speak. (p. 6)

Although Floridi addresses and emphasizes artificial intelligence, note that he casts it in the context of the digital—the implication being that AI is coalesced with other aspects of information communication technologies and draws upon them for its training protocols. Although it may appear as if one interacts uniquely with an AI platform, i.e., GPT Plus, there is an entire digital network in the background.

Some digital and AI decouplings:

1. intelligence and agency
2. location and presence
3. creation and authorship
4. diagnosis and expertise
5. experience and expertise

Specifically in education:
1. location and learning
2. time and learning
3. instructor and content delivery
4. libraries and information access
5. student engagement and instructor involvement

Some digital and AI decouplings:
1. diagnosis and expertise
2. algorithms and action
3. virtual and physical environments
4. pattern recognition and discovery
5. data analysis and human bias

Specifically in education:
1. instructor and student feedback
2. assessment and real time feedback
3. faculty development and teaching innovation
4. research collaboration across disciplines
5. curriculum design and skill development

Metaphorically, AI has “split the education atom” and shown that long held assumptions and actions can be separated or combined and in doing so may or may not improve teaching and learning. We may be sure, however, that the coupling or decoupling of these familiar structures will be disruptive. Moreover, in a complex system such as education, it is impossible to anticipate how an intervention will ripple through it. Often the outcomes will be counterintuitive and there will be unanticipated positive and negative side effects, all of which must be accommodated.

The Adjacent Possible for Education

The impact of artificial intelligence on education is undeniable but prediction for the long term becomes difficult because a proper baseline is so elusive (Rosling et al., 2018; Johnson, 2014; Roberts, 2007). Consider the protagonist Clay’s response to predicting the future in the book *Mr. Penumbra’s 24-Hour Bookstore* (Sloan, 2013):

> World government…no cancer…hover-boards.
> Go further. What’s the good future after that?
> Spaceships. Party on Mars.
> Further.
> *Star Trek*. Transporters. You can go anywhere.

> Further…
> I pause a moment, then realize: I can't. We probably just imagine things based on what we already know, and we run out of analogies. (p. 60)

However, there are foreseeable near-term implications for artificial intelligence in contemporary education. Driven by AI, customized learning paths can be integrated into the curriculum. This can be augmented with increased accessibility and interactive learning. There are other possibilities, such as better student and faculty collaboration or retention support. In the larger context, large language models will offer a wide range of educational assets that can relieve teachers and students from routine educational tasks and support more effective educational systems, while accommodating the necessary responses to societal changes.

At the same time, education will face AI challenges. One of the most concerning is the overreliance on platforms producing metacognitive laziness—reducing insight, creativity, and curiosity. The effect of something like Maxwell's "demon" (Maxwell, 1871) might convince students that knowledge can be gained and learning accomplished with no work involved. Just as physicists dispatched the demon by showing that gaining information involves work, the same is true for learning. There are other challenges—for instance, inherent bias in platforms acquired during their training protocols. Dickens was foreshadowing AI when he described a period of revolution as the best and worst of times (Dickens, 1859). For instance, our new understanding of interaction will be that of continually reframing our prompt queries because of the misalignment between what we thought we asked, and what AI responded to because it lacks contextual understanding.

Einstein and Carroll: Special Relativity and Individually Paced Learning

Similar principles apply in physics and education. In 1905, Albert Einstein's monumental thought experiment, special relativity, showed that the speed of light in all frames of reference under uniform motion is constant (Stoddard, 2023). Therefore, because of this constancy, the observed time must expand or contract. It dilates for speeds relative to the observer—meaning that for different reference positions time must be experienced differently to reflect light's constant speed (French, 1968; Billstrom, 2024). In 1963, John Carroll applied the same principle to education (Carroll, 1963). He showed that if learning is constant, then the time required must be flexible just as time is in special relativity. Students learn at different rates, so rather than place them in fixed time circumstances, such as semesters, students that require more time should be given it and those needing less should be accommodated. Again, learning time varies from the reference frame of the observer—student or instructor. Both Einstein and Carroll, in completely different contexts, arrived at the same conceptual conclusion that when one aspect of an environment is constant (light in physics and learning in education), another, such as time, must compensate. The principle supports adaptable approaches to education where instructional time flexes to ensure student mastery just as time bends in special relativity to accommodate light. In the education space, AI is likely to play a key role in bending learning time.

The Disappearing First Rung: How the Answer Machine Will Replace Entry Level Jobs and Impact Education and Organizations

Generative AI and related modern technologies will significantly change and in many cases replace entry-level jobs for higher education graduates, removing a foundational rung in new graduates' career development. Understanding how this shift will impact graduates, organizations, and education alike is essential to mitigate the risks and leverage the opportunities it creates for all three groups.

The chapter "Sprint to 2027" in Moskal et al. (2023), describes how one technology company is preparing for the shifting speed and scale modern technologies brings to its organizational culture. These rapid developments create an urgent need for a strategy to upskill the existing workforce while reshaping the methods and processes employees use to learn, think, and problem-solve. That argument is taken at least one step further given how Generative AI has dramatically escalated the technological impact on the workforce. The implications AI has for entry-level jobs that are traditionally the training ground for new graduates learning how to think in real-world environments creates a disappearing first-level career ladder rung. These roles, although often not glamorous, were formative. However, by 2030 a substantial share of roles in areas such as data entry, customer service, telemarketing, and financial and business analysis will be significantly reshaped by automation of routine tasks and AI-augmented work (World Economic Forum, 2023; McKinsey Global Institute, 2023; Forrester Research, 2024). In fact, it has already started: by late 2025 overall hiring had weakened. This weakening was mostly in tech and knowledge-work roles such as marketing and human resources, but job postings explicitly mentioning AI continued to grow, likely signaling a reallocation of hiring toward AI-related capabilities rather than broad employment expansion (Indeed Hiring Lab, 2026). The short-term cost efficiencies of Gen AI technologies are too attractive for commercial enterprises to resist even if they negatively impact long-term talent development.

The decrease in entry level jobs means new graduates will have fewer opportunities to start in roles where they confront ambiguity, gather information, and observe how decisions are made in interpersonal settings. Unfortunately, these experiential knowledge and skill sets are not readily acquired in classrooms, and increasingly less so post-COVID (Legon et al., 2023). However, if educational institutions collaborate with employers to reimagine workforce preparation where employees are fluent in enabling and leveraging AI, new graduates will be better prepared by being able to apply more critical thinking skills and fill higher-functioning roles beyond the traditional entry-level first rung positions.

Workplace Impact: The First Rung Will Disappear

In the traditional organizational model, new entry level employees with higher education degrees used their roles to gain fluency in complex, often ambiguous work environments. These proving grounds were where they gained, practiced, and refined critical thinking skills, eventually developing competence. However, increasingly, students use the generative AI embedded in standard software like Microsoft Office apps. While AI use has dramatically improved the polish and style of students 'output, it often belies a cognitive disengagement from the subject matter, and frequently results in writing that is generic, hackneyed, and often easily recognizable as AI slop. Put another way, they understand the form of what is being written and some of the context while expecting the polish AI provides, but they do not deeply understand the underlying concepts. In a professional development context, this is a missed opportunity to develop a deeper understanding of the subject area in which they are working and develop the employee´s individual style and persona.

Hard to Find Skills and Economic Drivers

Consistently over the past decade, in North America and Europe, employers reported in surveys that critical thinking, problem-solving, and analytical reasoning are among the most highly valued skills, they are also the most frequently lacking in recent graduates. In a 2020 AAC&U survey, sixty percent of employers said critical thinking was “very important,” yet only thirty nine percent believed graduates were very well prepared. That twenty-one-point gap has been stubbornly consistent. Put simply, employers want people who can question, synthesize, evaluate and who can make sense out of ambiguity. But if higher education institutions are only modestly successful at developing these skills, what happens when the early career roles that functioned as training grounds are themselves automated?

Strong Economic Incentive for AI

According to McKinsey (2023), generative AI could add up to $4.4 trillion annually in economic value across sectors by making processes less capital intensive and more cost efficient. In both manufacturing and service industries, adopting AI is already showing measurable cost reductions. Almost thirty percent of the business leaders McKinsey surveyed in 2024 reported using AI to directly cut costs. From customer support to data analysis and other entry-level functions AI systems can now perform work that used to be assigned to new graduates—often faster, cheaper, and without the cost of onboarding employees while leveraging the economies of scale AI technology provides. This rational economic behavior is reshaping how and when people learn to think professionally. The risk is that a generation of graduates will find themselves not only displaced from these roles but also underprepared for the kinds of judgment intensive work AI still cannot do while employers will not have the supply of human talent needed.

What is often missing from this discussion is a cost-benefit analysis of what organizations lose by eliminating those early-career positions. These first rung jobs are pathways to long-term talent development and part of how organizational knowledge, values, and culture are passed on. Without a first rung, organizations risk undermining their own talent pipeline, especially in roles that require collaboration and ethical and contextual decision-making. Additionally, there is a secondary motivation for enterprises and even governments to delegate decisions to algorithms and systems so that the institution may be judged more leniently in the event of a bad outcome (Feier et al., 2021).

While workplaces are replacing entry-level roles with AI, educational institutions are grappling with AI´s impact on learning.

Learning with the Answer Machine—Educational Institutions

In parallel with workplace adoption, students are increasingly using generative AI tools like ChatGPT to complete their academic work. The research on what this does to learning outcomes is still emerging, but increasingly troubling patterns are becoming more visible. A randomized controlled study from Peking and Monash Universities (Fan et al., 2024) found that while ChatGPT helped students improve essay scores in the short term, it did not increase knowledge transfer or intrinsic motivation. More concerningly, it triggered a form of what the researchers call "metacognitive laziness"—a diminished ability to monitor, evaluate, and adjust their thinking processes when AI is always ready to suggest the next move (Fan et al., 2024). This aligns with earlier research on "cognitive offloading," where the brain delegates problem-solving to external tools. That can be helpful when navigating complexity but damaging if it becomes habitual (Risko & Gilbert, 2016). As psychologist Adam Alter put it: "Disfluency, the sense that something is hard, is often the trigger for analytical thinking" (Alter et al., 2007). When that friction is removed, so is the cognitive workout that builds real capability. Similar dynamics in studies on navigation have found that when people rely solely on GPS or digital aids, their ability to form mental maps and understand spatial relationships diminishes. In one study, users of digital navigation aid performed worse on wayfinding tasks and spatial recall than those who navigated unaided (Parush et al., 2007). Technology got them to their destination, but it did not teach them how to get there.

Prohibiting AI Is Not the Answer

Institutions respond by prohibiting and trying to detect AI-generated text. This response will fail for the same reasons prohibiting calculators from primary level education was unsuccessful. Arguably, it is more beneficial for students to learn mathematical reasoning rather than calculation skills (University of Oxford, 2012). Finding the right balance is a lesson learned from calculator prohibition experience. In any case, AI detection tools are the lie detectors of our time. AI detection software is documented as unreliable, prone to false positives and fundamentally reactive (Yuan et al., 2023). More importantly, AI detection tools do not address the core issues of learning. Students are using AI not only to cheat, but to cope with performance expectations, information overload, the pressure to do research quickly and the need to feel competent. Even more troubling is recent research showing that one third of teenagers are using Generative AI for companionship that is very sycophantic (constant validation, agreement) in a way that the authors posit stunt emotional regulation that does not allow teens to learn to think critically when someone questions their assumptions (Robb et al., 2025).

If Gen AI coping mechanisms undermine long-term mental and emotional growth, the solution is not punitive detection but a reframing of AI use much like how calculators were eventually integrated into education. Concurrently, policymakers must enact age-appropriate guardrails in primary education that protect them from technology that inhibits the development of their cognitive abilities and emotional coping skills that are foundational to success in the work environment.

Reframing AI in Learning

The question is not how to stop students from using AI, it is how to change education so that AI becomes part of and improves the learning process rather than a substitute for it. There are increasing examples of this reported in the press (Silber, 2024) as well as in cases where when resources are applied AI improves outcomes (Powell, 2024). Using AI to improve outcomes requires a reframing as a thinking partner or tutor rather than an "answer machine." Scholars like Holmes et al. (2022) have begun recasting AI as an augmentative tool that can support metacognition if used reflectively. This requires reframing the "answer machine" to being a Socratic dialog that encourages self-reflection.

To achieve this, educational institutions should consider:

1. Reward reasoning and problem-solving more than output.
2. Require reflection on the use of AI tools—how they help, what they and students miss when using them.
3. Teach students to interrogate AI content for bias, incompleteness, and oversimplification.
4. Shift grading away from polished correctness and toward documented process.
5. Implement age-appropriate guardrails to ensure cognitive and emotional development occurs optimally.

Without this shift, we risk graduating students who can prompt AI but not think beyond it.

Where to Start

Given the pace of AI development, there is a narrow window in which to respond. Employers will need to articulate the skills they need, and how they will help develop them in a changing work structure. Universities will need to move beyond content delivery towards intentional incubators of cognition. And educators must resist the temptation to see AI only as a threat or a tool in a binary sense; it is a medium, and like any medium, it shapes how we think. How can employers and universities share responsibility for redesigning this new learning-to-work transition? The answer is not clear, but more conversation, research, sharing of needs and practical approaches to developing skills are good first steps.

Balanced approaches are emerging across educational institutions in Europe and the United States. Current research highlights both the potential benefits and challenges of generative AI, while simultaneously outlining balanced strategies for progress (Vieriu et al., 2025)

Perhaps AI skills will develop naturally, and students will focus on content more than form (e.g., skipping menial work) and critically engage with context. The answer machine will keep improving and doing more than humans. The analogy to tellers replaced by ATMs and internet banking is obvious but AI will do more. However, unless we deliberately intervene soon it will do more than replace the first rung on the career ladder. It will hollow out the human readiness those jobs once cultivated for a coming generation and over time it will erode the human capital that enterprises need to be profitable.

AI and The Arts of Inquiry

> The will to learn is an intrinsic motive, one that finds both its source and its reward in its own exercise. The will to learn becomes a 'problem' only under specialized circumstances like those of a school, where a curriculum is set, students confined, and a path fixed. The problem exists not so much in learning itself, but in the fact that what the school imposes often fails to enlist the natural energies that sustain spontaneous learning—curiosity, a desire for competence, inspiration to emulate a model, and a deep sense of commitment to the web of social reciprocity. Our concern has been with how those energies may be cultivated in support of school learning. (Bruner, 1966)

> The condition for understanding the Truth is like the capacity to inquire for it. (Kierkegaard, 1846)

A Mind at Work

Vannevar Bush's seminal essay "As We May Think" (1945) identifies two enormous difficulties faced by scientists: knowledge that is obscure or cannot be found, and knowledge that cannot be put to use because the pace of discovery outruns our minds.

AI can help with both of those problems. When Bush describes the "growing mountain of research" that leaves investigators "staggered by the findings and conclusions of thousands of other workers … conclusions he cannot find time to grasp, much less to remember, as they appear," we can gratefully turn to AI summaries as one possible solution. When Bush analyzes "methods of transmitting and reviewing the results of research" that are "totally inadequate for their purpose," he offers the chilling example of Gregor Mendel's genetics research: "lost to the world for a generation because his publication did not reach the few who were capable of grasping and extending it" (p. 37). Here AI-enhanced searching brings hope.

Yet, as Bush describes the need to avoid "truly significant attainments becom[ing] lost in the mass of the inconsequential" (p. 37), we should recall that s*ignificant* and *inconsequential* are value judgments best crafted not *for* our species but *by* our species, the species that searches for meaning. AI has the unwelcome potential to ease the burden of discernment onto its stout and confident mechanical shoulders, with the disastrous result that human beings will accede to a world in which their search for meaning becomes merely an acceptance of proliferating algorithmic judgments delivered by convincingly modulated pseudo-human voices, both oral and written. And we won't know it's happening, or even that it has happened.

Automatic machine-generated judgments, however, are only part of the risk our species faces with these new technologies. Perhaps the greatest risk of all is the loss of the arts of inquiry. How do we imagine the role of a question, and in turn, the character of an answer? Data may be served up like cash from an ATM, but the search for insight and fresh understandings depends on far more complex, provisional, and courageous modes of inquiry. As various kinds of AI increasingly constitute not only popular culture but the culture of education itself, the stamina and imagination such modes of inquiry require may recede into the past and ultimately be lost.

For Bush's vision in "As We May Think" is not limited to findability and curation. His vision is also one of sharing intellectual *exploration*. His focus at the climax of the essay shifts to collaborative knowledge generation fostered by shared histories of investigation. Bush's famous "Memex" was not just a "memory extender" but a means of recording and sharing what Bush called the "associative trails" marked by embodied minds as their inquiries take them hither and yon to the novel insights their wandering, varied lives empowered. Indeed, Bush called for "a new profession of trail blazers, those who find delight in the task of establishing useful trails through the enormous maze of the common record. The inheritance from the master becomes, not only his additions to the world's record, but for his disciples the entire scaffolding by which they were erected" (p. 46). These trails are not formulae or algorithms, repeating the same set of approaches uninflected by fresh experience (machines do not have experiences, merely reactions), but the paths of thought, imagination, and idiosyncratic associations that together constitute a master class in the arts of inquiry. The trail blazer, in sum, "builds *a trail of his interest* through the maze of materials available to him" (p. 46, emphasis added).

Bush's Memex thus becomes a moving picture of the mind at work in unpredictable ways that, inflected by personal experience and a stockpile of proximal and potential connections, may break through the "already" into "not yet." Anyone blessed with a great teacher recognizes that an essential part of learning is attending to, and mulling over, the drama of inquiry present in the teacher's articulation of knowledge, a drama all the more powerful for being suggestive rather than prescriptive. This drama of inquiry is what Michael Nielsen (2012) describes in *Reinventing Discovery: The New Era of Networked Science*: "It's not just the scientific content that matters, it's the culture that is revealed, a particular way of viewing the world…. [T]he first time in my life that I heard a scientist speaking informally was when I was 16. It changed my life" (p. 168–169).

Integrated Domains and Oblique Strategies

One of the most far-seeing prophets of the digital age in this regard was Douglas Carl Engelbart. Most famous as the inventor of the computer mouse, Engelbart's accomplishments far outstrip that innovation, extending to a comprehensive vision of enhanced communication and collaboration in complexly linked networks. In his 1962 manifesto *Augmenting Human Intellect: A Conceptual Framework*, Engelbart prefaces his long and dazzling argument with this challenging statement:

> We do not speak of isolated clever tricks that help in particular situations. We refer to a way of life in an integrated domain where hunches, cut-and-try, intangibles, and the human "feel for a situation" usefully co-exist with powerful concepts, streamlined terminology and notation, sophisticated methods, and high-powered electronic aids. (p. 1)

Engelbart understood, as only a small handful of people did at the time, that networked digital computing represented a fundamental change in the way human beings represented and communicated about their experience of the world. As Clay Shirky (2008) put it, some forty-six years later, "any radical change in our ability to communicate with one another changes society" (p. 106). Engelbart could see that these changes could bring unprecedented benefit, but he also knew that computing could fall far short of its potential benefit or even become quite harmful. Engelbart understood that our ingenuity could become a trap baited with clever tricks, a situation in which, having been shaped by our tools, we lost our power to shape those tools.

For the arts of inquiry do not proceed along linear trajectories. They do not lend themselves to algorithmic formulation. The arts of inquiry rely on the human ability to think abstractly and by analogy—analogy being, as Douglas Hofstadter (2001) insists, the core of cognition (p. 499–500). The arts of inquiry must be elicited, suggested, modelled, encouraged, not mapped or outlined or simply "transmitted." They are stimulated, for example, by exercises such as Brian Eno's and Peter Schmidt's "Oblique strategies: Over one hundred worthwhile dilemmas" (1975), a set of cards in which each card conveys an instruction that by design is *not* directly related to the task at hand. Yet these instructions are not irrelevancies, but wise advice emerging from the creative experiences of Eno and Schmidt themselves:

> These cards evolved from separate observations of the principles underlying what we were doing. Sometimes they were recognized in retrospect (intellect catching up with intuition), sometimes they were identified as they were happening, sometimes they were formulated.
>
> They can be used as a pack (a set of possibilities being continuously reviewed in the mind) or by drawing a single card from the shuffled pack when a dilemma occurs in a working situation. In this case the card is trusted even if its appropriateness is quite unclear. They are not final, as new ideas will present themselves, and others will become self-evident.

Such oblique strategies strongly resemble what Doug Engelbart (2003) wrote forty years after *Augmenting Human Intellect* as he considered the uniquely important goal of not simply improving a process but "improving our ability to improve," an activity he termed a "C-level" activity, where "A" is the initial process or goal and "B" is the process of improving "A." Engelbart recognized that without that "C" level, humankind cannot "bootstrap" itself into new ways of thinking, but will be chained to the endless task of refining current practices instead of reinventing them in a true advance.

> At the C level we are trying to understand how improvement really happens, so that we can improve our ability to improve. This means having different groups exploring different paths to the same goal. As they explore, they constantly exchange information about what they are learning. The goal is to maximise [sic] overall progress by exchanging important information as the different groups proceed. What this means, in practice, is that the dialog between the people working toward pursuit of the goal is often just as important as the end result of the research. Often, it is what the team learns in the course of the exploration that ultimately opens up breakthrough results. (p. 162)

To what extent does the culture of AI stimulate exploration, and to what extent does it substitute spoon-feeding for exploration even as it conceals that substitution behind an apparently information-rich and confidently authoritative "voice"?

Engelbart (2003) goes on:

> [A]t the C level … [w]e are not trying to solve a specific problem, but … reaching for insight into a broad class of activities and opportunities for improvement…. I have routinely found that when I seem to reach a dead end in my pursuit of a problem, the key is usually to move up a level of abstraction, to look at the more general case…. [T]his is directly counter to the typical approach to solving focused, B-level problems, where you typically keep narrowing the problem in order to make it more tractable. In our work on improving improvement, the breakthroughs come from the other direction—from taking on an even bigger problem. (p. 162–163)

I can imagine forms of AI prompting its users to move up a level of abstraction, to adopt oblique strategies. Yet this would mean reimagining our relationship not only to AI, but to learning itself, especially within the culture of education. My fear is that AI, of whatever variety, will seduce us into focusing only on the tractable, and exclude the very notion of wicked problems or open-ended searching. Our ability to get to what Engelbart terms the "C-level," another name for the arts of inquiry, will weaken and die.

Inquiry and Wisdom

The phrase "arts of inquiry" may seem too abstract in the current heated conversation about practical benefits conferred by generative and agentic AI, so let's consider two specific risks: the diminishment of insight and the diminishment of wisdom. In contrast to knowledge, insight results from new neural connections, an "a-ha!" experience familiar to teachers and learners and a neuropsychological phenomenon that can be measured by EEGs (Kandel, 2012). Eric Kandel describes psychologist Jonathan Schooler's conclusions in this regard:

> All of a sudden, ideas that were previously isolated come together and people see connections that had escaped them—and others—before…. [P]roblems requiring creative insight are very likely to lead to an impasse—a state in which the person does not know what steps to take next; and … such problems are quite likely to lead to a sustained effort that is rewarded with a sudden insight that breaks the impasse and clearly reveals the solution. This situation is thought to arise when the problem solver breaks free of unwarranted assumptions and forms novel, task-related connections between existing concepts or skills. (p. 458)

Kandel concludes that "[a]lthough problem solving generally relies on shared cortical networks, the sudden flash of insight that occurs when we engage distinct neuronal and cognitive processes seems to allow us to see connections that had previously been elusive" (p. 482).

What becomes of the potential for such connections—such insights—when learners rely on Gen AI to write their papers for them? A study by MIT researchers (Kosmyna et al., 2025), not yet peer-reviewed at the time of this writing but available as a preprint from arXiv.org, found that using ChatGPT to write an essay substantially reduced intraneural connectivity, as well as diminishing the ability to feel "ownership" of one's work and quote from it reliably and extensively. Moreover, the disposition of students using genAI extensively and indiscriminately is chillingly demonstrated by a *New Yorker* article featuring "Alex" (a pseudonym), a student at New York University who laughingly admits to the truth about the two genAI-written essays he had turned in during finals: "I didn't retain anything…. I couldn't tell you the thesis for either paper hahhahaha" (Hsu, 2025).

This is not to say that no legitimate, cognition-strengthening uses of generative and agentic AI exist. The *New Yorker* article cited above points out some of them, and every day the news from the sciences suggests both startling potential and realized advances in laboratory analysis, medical diagnosis, and other essential areas. Nevertheless, the MIT study (Kosmyna et al., 2025) sounds a clear and urgent warning we do well to heed: "in this study we demonstrate the pressing matter of a likely decrease in learning skills…. The use of LLM[s] had a measurable impact on participants, and while the benefits were initially apparent, as we demonstrated over the course of 4 months, the LLM group's participants performed worse than their counterparts in the Brain-only group at all levels…" (p. 2). The researchers conclude that "longitudinal studies are needed in order to understand the long-term impact of the LLMs on the human brain, before LLMs are recognized as something that is net positive for the humans" (p. 143).

And what of wisdom, which one might call the accumulation of insight across a lifetime? Here we might turn to the research on wisdom undertaken by investigators across disciplines and at centers ranging from the Center for Lifespan Psychology at the Max Planck Institute for Human Development (https://www.mpib-berlin.mpg.de/research/research-centers/lifespan-psychology) to the Center for Practical Wisdom at the University of Chicago (https://wisdomcenter.uchicago.edu/). One influential theory of wisdom, the Berlin Wisdom Paradigm, connects knowledge to wisdom with the concept of "wisdom-related knowledge" assessed on the basis of "rich factual knowledge, rich procedural knowledge, life-span contextualism, value relativism, and uncertainty" (Steinberg and Jordan, 2005, p. xiii). Even a more cautious approach toward assessing wisdom, such as seen in "Can we measure practical wisdom?" (Swartwood, 2020), concludes that the goal is worthwhile and requires an essentially interdisciplinary approach:

> Philosophical reasoning can help us specify which practical wisdom-relevant characteristics (e.g., reasoning or discrimination skills or processes, self-regulation skills or habits) are logically entailed by a rationally defensible account (such as the minimal philosophical conception). Empirical research can then determine how these can be developed and how they relate to other characteristics whose necessity for practical wisdom is more contentious or contingent on facts about human psychology. (For example, does an ability to give advice or justify one's choices tend to develop as a result of acquiring the reasoning and self-regulation processes that are part of a philosophically plausible picture of wisdom?) Even if the results of such an investigation would not be sufficient to allow us to measure practical wisdom or its components, they still would provide valuable information about some of the necessary conditions for having or developing practical wisdom and a general picture of some of the ways it would be likely to manifest in real human lives. (Compare general best practice measures of teaching: they can help us acquire a fuller picture of good teaching and how we can improve, even if they do not constitute measures of teaching expertise and should not be used as such.) (p. 91)

How might generative or agentic AI serve to develop and strengthen those essential elements of wisdom? It would be foolish (the opposite of wise) to argue such benefits are impossible. Yet it is fair to say that the multi-billion-dollar gold-rush mentality on display at the leading platforms for AI research—Google, Meta, Claude, OpenAI, etc.—does not suggest a thoughtful and usefully cautious approach to this question. Instead, as prior sections of this very article make clear, the potential damage wrought by indiscriminate uses of generative and agentic AI may pose existential threats to human culture and civilization: in short, a dark age, defined by Jane Jacobs (2004) as a time when even the memory of a more enlightened age disappears.

Indeed, in her final book, *Dark Age Ahead*, Jacobs describes a culture of education badly in need of reform:

> All universities possess their own subcultures, and so do departments within universities, varying to the point of being indifferent or even antagonistic to one another, so a generalization cannot describe all accurately. But it is safe to say that credentialing as [the] primary business of institutions of higher learning had gotten under way in the 1960s. Students were the first to notice the change. In the unrest and turbulence of that decade, one thread of complaint came from students who claimed they were shortchanged in education. They had expected more personal rapport with teachers who had become only remote figures in large, impersonal lecture halls. The students were protesting attempts to transmit culture that omitted acquaintance with personal examples and failed to place them on speaking terms with wisdom. In another decade, however, students dropped that cause, apparently taking it for granted that credentialing is the normal primary business of institutions of higher learning and that its cost is an unavoidable initiation fee into acceptable adulthood. (p. 46–47)

The goal of being on "speaking terms with wisdom," always ambitious, may recede forever out of reach without a much more thoughtful and discriminating exploration of AI. If personal rapport with teachers and other experiences of the arts of inquiry become mere "prompt engineering," what becomes of insight and wisdom? When teaching assistants are agentic AI bots, a program the University of Michigan and Google are already implementing ("Google Public Sector," 2025), when premium genAI corporations offer their products for free to students, with a wink, saying "Good luck with finals!" (Shroff, 2025), lifelong learning degenerates into brand loyalty and individual agency becomes an illusion. Higher learning partnering with a global data-hoovering behemoth of unparalleled wealth and influence. Students having the very best cheating resources freely available at a critical moment in strengthening and assessing their learning. What could go wrong? The profit from learning used to be metaphorical and applied to the learners themselves. Now the profits accrue to the corporations as non-profits brag about student success and for-profits brag about market penetration and shareholder returns. Do such strategies nurture the arts of inquiry, foster insight, or inculcate wisdom? Do they reveal a commitment to nurturing "creative, curious, caring, and collaborative learners," the "4C" approach advocated by MIT Media Lab Professor of Learning Research Mitchel Resnick (2025)? I am skeptical.

It need not be this way. Perhaps there is still time to pursue the arts of inquiry in the culture of education. The beneficial potential of AI can be realized only in such a pursuit. Perhaps we can rediscover and renew our commitment to the idea of school as what Bruner calls "an exercise in consciousness-raising about the possibilities of communal mental activity" and not simply a transactional environment narrowly imagined as "a means for acquiring knowledge and skill" *(Culture of Education,* 1996). Indeed, we must. The alternative is dire: a self-satisfied, AI-generated somnambulism from which we may never awaken.

Finally: A Way Forward

The concept and functionality of artificial intelligence struggled for decades because of computational limitations, data scarcity, misunderstanding intelligence, and reliance on symbolic approaches. This led to multiple cycles of overpromising and underdelivering. However, in recent years advances in deep learning, big data coupled with computational sophistication, and hardware make it clear that artificial intelligence, digital education, information availability, organizational onboarding, human thinking and creativity, quality, and the roles of teachers and students will be impacted. AI has resulted in positive outcomes. In medicine, more effective diagnostic and treatment procedures are being developed daily. Self-driving vehicles and traffic management are reframing transportation. Newly developed assistive AI technologies are improving the lives of people with disabilities. Accelerated data and text analysis, predictive modeling, and simulation techniques are accelerating scientific research. Cybersecurity has increased. Real-time language translation is creating effective communication patterns. In education, curriculum development is progressing. Teachers are being relieved of time consuming non-instructional tasks. Professional development is improving. Coaching and mentoring services powered by AI increase students' learning effectiveness. AI has taken adaptive learning to a whole new level of effectiveness. Big data analytics and pattern recognition have accomplished information processing results heretofore thought impossible. Learning environments are being optimized by intelligent management systems. The opportunities for diverse learners have increased. These comprise a small sample of how artificial intelligence can improve the human condition. However, as we attempt to frame the context of how AI should be considered it is important not to become binary in our thinking—as in AI is good or bad. The question is more nuanced. We must build the proper reference contexts so that we can better understand the long-term impacts of modern technologies that can create things and ideas, make decisions, and function as autonomous agents that we may not be able to control.

The potential for humankind is limitless but the ways forward are impossible to understand at present. The rapidity with which this is happening surpasses most everything we have experienced in the past. However, there are consequences on the horizon and the future is uncertain at best. By what baselines are we to judge these developments? What if we have entered an era of hyper history where information communication technologies including artificial intelligence communicate only with each other, leaving humans out of the loop? What are the implications for transcending their training and creating, conceptualizing, and communicating in a nuanced and symbolic manner that we may not understand? This may seem like science fiction, but it is a possibility. As mentioned in this article, it is impossible to sense how these developments will impact our lives, and clearly, we will be challenged by unanticipated side effects. As a result, we find ourselves caught up in a web of ambivalence: enthusiastic anticipation and simultaneous concern and foreboding. So many possibilities for enabling humankind are on the horizon. For instance, consider the potential for learning. The fundamental components of contemporary learning theory are:

1. Behaviorism focuses on observable behavior shaped by instructional environments where behavior change indexes learning.
2. Cognitivism emphasizes mental processes such as memory, attention, and problem solving where knowledge acquisition reflects learning.
3. Constructionism where students formulate their unique understanding through experience and interaction.
4. Connectivism depends on the digital network where learning resides in a network structure.

We can hypothesize that artificial intelligence will reframe those learning components into an enhanced and more comprehensive theory. For example, knowledge might be viewed as a dynamic network based on emergent curricula. The instructor might gravitate further to the facilitator and ethical counselor role. The learner will abandon the remnants of passive learning, reconfigure cognitively, and become more self-directed while maintaining human analytic capabilities. The learning environment will embrace ambient intelligence, simulation, and coalesced formal and informal learning. A reasonable assumption is that artificial and human intelligence will interact in a yet unknown complex system where the learning theory of the future will depend more on the elements' interactions instead of their definitions. The long-term and unanswered question is, however, will a meaningful balance between human and artificial learning result?

One of the fundamental problems in AI is alignment, that is, ethical guidelines and guardrails for AI that help ensure humane outcomes in support of human flourishing. Ironically, alignment depends on just the kinds of learning in practical wisdom that widespread use of AI in education, by both students and teachers, may erode or irreparably damage. In an interview on the podcast *The Good Fight with Yascha Mounk*, technical advisor Dean Ball describes the nature and importance of the alignment problem: "People should not dismiss the notion of alignment. It is a real thing. But they assume that it is a problem susceptible to a solution, like one magic solution, like an equation that can be solved and is one hundred percent solved at that point. We do not think alignment is that. We think alignment is much more of what we would describe as a muddle-through problem. Alignment is much closer to the articulation of philosophy.... [I]f the alignment problem were susceptible to a one hundred percent solution, that would be equivalent to solving philosophy.... {But if] doing philosophy well is an important part of alignment, then that is something we know how to do. We are not perfect at it, but it is not something you can ever perfect. We have practice, experience, and a deep body of knowledge to draw on.... Perhaps we are being too optimistic. We do not want to leave the impression that the work is done or that this model [Anthropic's "Soul Document"] solves alignment. It does not...." (Mounk, 2026).

But alternatively, threats to integrity and our value systems loom on that same horizon. AI systems are becoming increasingly opaque and less transparent. Mathematically complex components such as tokens defy direct understanding and become the object of awkward metaphors. Knowledge acquisition becomes a series of perturbations with these systems. Therefore, we leave the readers with a series of long-term questions for their consideration.

1. How can educational systems be designed to maximize AI's potential for enhancing cognitive skills while strengthening analogous abstract and nuanced problem solving?
2. How will educators and students accommodate an increasingly autodidactic educational environment?
3. How can we ensure recency of AI-generated information for critical applications like education when these platforms are frozen at their time of training?
4. How can we ensure that AI-driven personal benefits exist for learners while excluding existing inequalities in training data?
5. What is the emerging paradigm of authorship, authority, and intellectual property?
6. Considering the power law distribution of information sources, how will we develop information fluency skills in students and what role will educators play in the process?

7. How can we incorporate AI systems and educational theories and practices that promote active learning and authentic understanding rather than passive acquisition of information—is AI unraveling academia?
8. Is it possible for AI to become a partner rather than a dominance in the inquiry process, and if so, how will this change the pursuit of knowledge?
9. How can we foster genuine interdisciplinary collaboration, learning, and research in the transformed AI educational environment?
10. What integrity-review-and assessment processes will educators develop and adopt to screen out "bad actor" AI made possible by open source frontier AI, particularly when bad-actor interventions (one might say, "hacking") in training and alignment can have catastrophic results for learners' lives?
11. How can curricula re-emphasize core learning in the humanities, especially ethics and moral philosophy, to help students judge the character and merits of AI systems and reliably infer the alignments those systems have undergone?
12. How can we marshal a collective and thoughtful response to unavoidable and unanticipated side effects in large language models that produce inappropriate or harmful outcomes, such as the sycophancy problem, where AI disproportionately generates positive and validating responses to users regardless of their prompt accuracy or appropriateness?

Declarations

Supplementary Materials

There are no supplementary materials for this manuscript except for those provided by the references

Author Contributions

Each of the four contributed equally to this manuscript, which is signified by the fact that authorship is depicted in alphabetical order.

Funding

No funding from any source was received for this project.

Institutional Review Board Statement

There were no human subjects involved in the preparation of this manuscript therefore no data were collected nor analyzed.

Informed Consent Statement

Because no human subjects were involved in this manuscript preparation informed consent was not required.

Data Availability Statement

There were no data associated with this manuscript.

Acknowledgements

The authors acknowledge Drs. Anthony Picciano and Patsy Moskal, guest editors for this special edition, for inviting the submission of this manuscript.

Conflicts of Interest

There are no conflicts of interest associated with this project.

References

AAC&U. (2020). *How College Contributes to Workforce Success: Employer Views on What Matters Most*. https://www.aacu.org/research/how-college-contributes-to-workforce-success

AAC&U. (2025). *Higher education leaders navigate AI disruption*. https://www.aacu.org/newsroom/higher-education-leaders-navigate-ai-disruption

AAUP. (2025). *Faculty report declining academic freedom*. https://www.aaup.org/news/faculty-report-declining-academic-freedom

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Meta cognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*(4), 569–576.

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. Hyperion.

Banerjee, S., Agarwal, A., & Singla, S. (2024). *LLMs will always hallucinate, and we need to live with this.* arXiv.org https://arxiv.org/abs/2409.05746v1

Barabási, A.-L. (2016). *Network science*. Cambridge University Press.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407–424. https://doi.org/10.1108/eb024320

Berti, L., Giorgi, F., & Kasneci, G. (n.d.). Emergent abilities in large language models: A survey. *arXiv.org.* https://arxiv.org/pdf/2503.05788

Billstrom, R. (2024). *Visualizing special relativity: An intuitive approach.*

Brontë, C. (1847). *Jane Eyre*. Smith, Elder & Co.

Bruner, J. (1966). *Toward a theory of instruction*. Harvard University Press.

Bruner, J. (1966). *The culture of education*. Harvard University Press.

Bush, V. (2003). As we may think. In N. Wardrip-Fruin (Ed.), *The new media reader* (pp. 37–47). MIT Press. (Original work published 1945)

Caldarelli, G. (2007). *Scale-free networks: Complex webs in nature and technology.* Oxford University Press.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64*(8), 1–9. https://doi.org/10.1177/016146816306400801

Dickens, C. (1859). *A tale of two cities.* Chapman and Hall.

Dubois, D. (2024). Paradoxes of generative AI and the simultaneous promise and threat to academic freedom. *AAUP Journal of Academic Freedom, 15.* https://www.aaup.org/JAF15/paradoxes-generative-ai

Dziuban, C., Moskal, P., Reiner, A., & Cohen, A. (2023). Student rating and course modalities: A small study in a large context. *Online Learning, 27*(3), 70–103. https://doi.org/10.24059/olj.v27i3.4053

Engelbart, D. (1962). *Augmenting human intellect: A conceptual framework*. https://www.dougengelbart.org/pubs/papers/scanned/Doug_Engelbart-AugmentingHumanIntellect.pdf

Englebart, D. (2003). *Improving our ability to improve: A call for investment in a new future*. In *Global knowledge renaissance: Selected papers from the World Library Summit 2002*. National Library Board Singapore, Times Books International.

Eno, B., & Schmidt, P. (2013). *Oblique strategies. Over one hundred worthwhile dilemmas* (6th ed.). Opal Ltd.

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen Y., Li X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology, 56(2), 489–530. https://doi.org/10.1111/bjet.13544

Feier, T., Gogoll, J., & Uhl, M. (2021). Hiding behind machines: When blame is shifted to artificial agents. *arXiv preprint arXiv:2101.11465*. https://doi.org/10.48550/arXiv.2101.11465

Fitzgerald, F. S. (1925). *The great Gatsby*. Charles Scribner's Sons.

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Floridi, L. (2019). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.

Forrester Research. (2024, June 6). *AI and automation will take 6% of U.S. jobs by 2030*. Forrester.

French, A. P. (1968). *Special relativity* (1st ed.). W. W. Norton & Company.

Gladwell, M. (2024). *Revenge of the tipping point: Overstories, superspreaders, and the rise of social engineering*. Little, Brown and Company.

Stephen M. Ross School of Business, University of Michigan. (2025, April 7). *Google Public Sector helps enhance learning at the University of Michigan with pioneering new agentic AI virtual teaching assistant*.. https://michiganross.umich.edu/news/google-public-sector-helps-enhance-learning-university-michigan-pioneering-new-agentic-ai

Guo, Y., Guo, M., Su, J., Yang, Z., Zhu, M., Li, H., Qiu, M., & Liu, S. S. (2024). *Bias in large language models: Origin, evaluation, and mitigation.* arXiv.org https://doi.org/10.48550/arxiv.2411.10915

Harari, Y. N. (2024). *Nexus: A brief history of information networks from the Stone Age to AI*. Harper.

Hoffer, E. (1951). *The true believer: Thoughts on the nature of mass movements.* Harper & Brothers.

Hofstadter, D. (2001). Epilogue: Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 499–538). The MIT Press.

Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial Intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Hsu, H. (2025) What happens after A.I. destroys college writing? *The New Yorker.* https://www.newyorker.com/magazine/2025/07/07/the-end-of-the-english-paper

Hüyük, A., Xu, X., Maasch, J., Nori, A. V., & Gonzalez, J. (2025). Reasoning elicitation in language models via counterfactual feedback. *OpenReview.* https://openreview.net/forum?id=VVixJ9QavY

Indeed Hiring Lab. (2026, January 22). *January 2026 U.S. labor market update: Jobs mentioning AI are growing amid broader hiring weakness*. https://www.hiringlab.org/2026/01/22/january-labor-market-update-jobs-mentioning-ai-are-growing-amid-broader-hiring-weakness/

Jacobs, J. (2004). *Dark Age Ahead*. Random House, First Vintage Books Edition. 2004.

Johnson, S. (2014). *How we got to now: Six innovations that made the modern world*. Riverhead Books.

Kandel, E. (2012). *The age of insight: The quest to understand the unconscious in art, mind, and brain, from Vienna 1900 to the present*. Random House.

Kierkegaard, S. (1947). Philosophical fragments. In R. Bretall (Ed.), *A Kierkegaard anthology*. Princeton University Press.

Kosmyna, N., et al. (2025). "Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task." arXiv.org https://arxiv.org/abs/2506.08872v1

Kuhn, T. S. (1962). *The structure of scientific revolutions* (3rd ed.). University of Chicago Press.

Legon, R., Fredericksen, E. E., Simunich, B., & Garrett, R. (2023). *CHLOE 8: Student demand moves higher ed toward a multi-modal future*. Quality Matters & Eduventures.

Liu, G. K.-M. (2023). *Perspectives on the social impacts of reinforcement learning with human feedback.* arXiv.org https://arxiv.org/abs/2303.02891

Manus.im. (2025). *Manus.* https://manus.im/

Maxwell, J. C. (1871). *Theory of heat*. Longmans.

McKinsey & Company. (2023). *The economic potential of generative AI: The next productivity frontier*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

McKinsey Global Institute. (2023, July 26). *Generative AI and the future of work in America*. McKinsey & Company.

Meade, C., & Jano, P. (2024). *Hallucinations in LLMs: Types, causes, and approaches for enhanced reliability*. ResearchGate. https://doi.org/10.13140/RG.2.2.12184.61445

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.

Morville, P. (2005). *Ambient findability: What we find changes who we become*. O'Reilly Media.

Mounk, Y. (2026, March 7). Dean Ball on who should control AI [Audio podcast episode]. In *The good fight with Yascha Mounk.* https://writing.yaschamounk.com/

Nielsen, M. (2012). *Reinventing discovery: The new era of networked science*. Princeton University Press.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*(5), 323–351. https://doi.org/10.1080/00107510500052444

OpenAI. (2025). *Introducing deep research*. https://openai.com/index/introducing-deep-research/

Ovetz, R. (2025). *Refusing AI in higher education.* AAUP. https://www.aaup.org/article/refusing-ai-higher-education

Page, S. E. (n.d.). *Understanding complexity* [Video lectures]. The Great Courses.

Parush, A., Ahuvia, S., & Erev, I. (2007). Degradation in spatial knowledge acquisition when using automatic navigation systems. *International Journal of Human-Computer Interaction, 23*(3), 257–274.

Pilehvar, M. T., & Camacho-Collados, J. (2021). *Embeddings in natural language processing*. SpringerLink. https://doi.org/10.1007-978-3-031-02177-0

Powell, A. (2024, September 18). Professor tailored AI tutor to physics course; engagement doubled. *Harvard Gazette*. https://news.harvard.edu/gazette/story/2024/09/professor-tailored-ai-tutor-to-physics-course-engagement-doubled/

Purdy, M. (2024). What is agentic AI, and how will it change work? *Harvard Business Review*. https://hbr.org/2024/12/what-is-agentic-ai-and-how-will-it-change-work

Resnick, M. (2025). In the age of AI, we need a human-centered society more than ever. *Medium*. https://mres.medium.com/in-the-age-of-ai-we-need-a-human-centered-society-more-than-ever-0be3449ad40e

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688.

Robb, M. B., & Mann, S. (2025). *Talk, trust, and trade-offs: How and why teens use AI companions*. Common Sense Media.

Roberts, C. (2007). *The unnatural history of the sea*. Island Press/Shearwater Books.

Rosling, H., Rosling, O., & Rönnlund, A. R. (2018). *Factfulness: Ten reasons we're wrong about the world—and why things are better than you think*. Flatiron Books.

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. Penguin Books.

Shroff, L. (2025). The Gen Z lifestyle subsidy. *The Atlantic*. https://www.theatlantic.com/technology/archive/2025/04/college-students-free-chatgpt/682532/

Silber, T. (2024, April 1). How students are using ChatGPT to cheat—and what schools are doing about it. *The Wall Street Journal*.

Sloan, R. (2013). *Mr. Penumbra's 24-Hour bookstore Kindle*. Picador/Farrar, Straus and Giroux.

Smith, M. J., & Dziuban, C. D. (2023). Sprint to 2027: Corporate analytics in the digital age. In P. D. Moskal, C. D. Dziuban, & A. G. Picciano (eds.), *Data analytics and adaptive learning: Research perspectives* (1st ed.). Routledge.

Sternberg, R., & Jordan, J. (Eds.). (2005). *A handbook of wisdom: Psychological perspectives*. Cambridge University Press.

Stoddard, J. (2023). *Relativity for beginners, the special and the general theory: The 3 bizarre discoveries you must know to master Einstein's relativity fast, revealed step-by-step*.

Swartwood, J. (2020). Can we measure practical wisdom? *Journal of Moral Education*, *49*(1), 71–97. doi: 10.1080/03057240.2019.1702933

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. Random House.

Taleb, N. N. (2012). *Antifragile: Things that gain from disorder*. Random House.

Taleb, N. N. (2018). *Skin in the game: Hidden asymmetries in daily life*. Random House.

Tolstoy, L. (1878). *Anna Karenina* (trans. C. Garnett). William Heinemann.

University of Oxford. (2012, December 13). *Banning calculators in primary schools a backward step, academics warn*. Oxford University Department of Education.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc. https://arxiv.org/abs/1706.03762

Vieriu, A. M., & Petrea, G. (2025). *The impact of Artificial Intelligence (AI) on students' academic developmen*t. *Education Sciences*, *15*(3), 343. https://doi.org/10.3390/educsci15030343

Wadekar, S. N., Chaurasia, A., Chadha, A., & Culurciello, E. (2024). *The evolution of multimodal model architectures.* arXiv.org https://arxiv.org/abs/2405.17927

Wang, C., Deng, Y., Lyu, Z., Zeng, L., He, J., Yan, S., & An, B. (2024). *Q: Improving multi-step reasoning for LLMs with deliberative planning*. arXiv.org https://arxiv.org/abs/2406.14283

World Economic Forum. (2023). *The future of jobs report 2023*. https://www.weforum.org/publications/the-future-of-jobs-report-2023/

Xia, Y., Kim, J., Chen, Y., Ye, H., Kundu, S., Hao, C., & Talati, N. (2024). *Understanding the performance and estimating the cost of LLM fine-tuning*. arXiv.org https://arxiv.org/abs/2408.04693v1

Yuan, X., Zhang, X., & Zhao, H. (2023). Detecting AI-generated text: Challenges, tools, and ethical concerns. *AI Ethics, 7*(1), 23–36.

Zhang, S., Chen, Z., Chen, S., Shen, Y., Sun, Z., & Gan, C. (2024). *Improving reinforcement learning from human feedback with efficient reward model ensemble.* arXiv.org https://arxiv.org/abs/2401.16635v1

Zhang, W., & Zhang, J. (2025). Hallucination mitigation for retrieval-augmented large language models: A review. *Mathematics, 13*(5), 856–856. https://doi.org/10.3390/math13050856

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.