

Opinion Short Paper

The Dawn of Intelligence Augmentation?

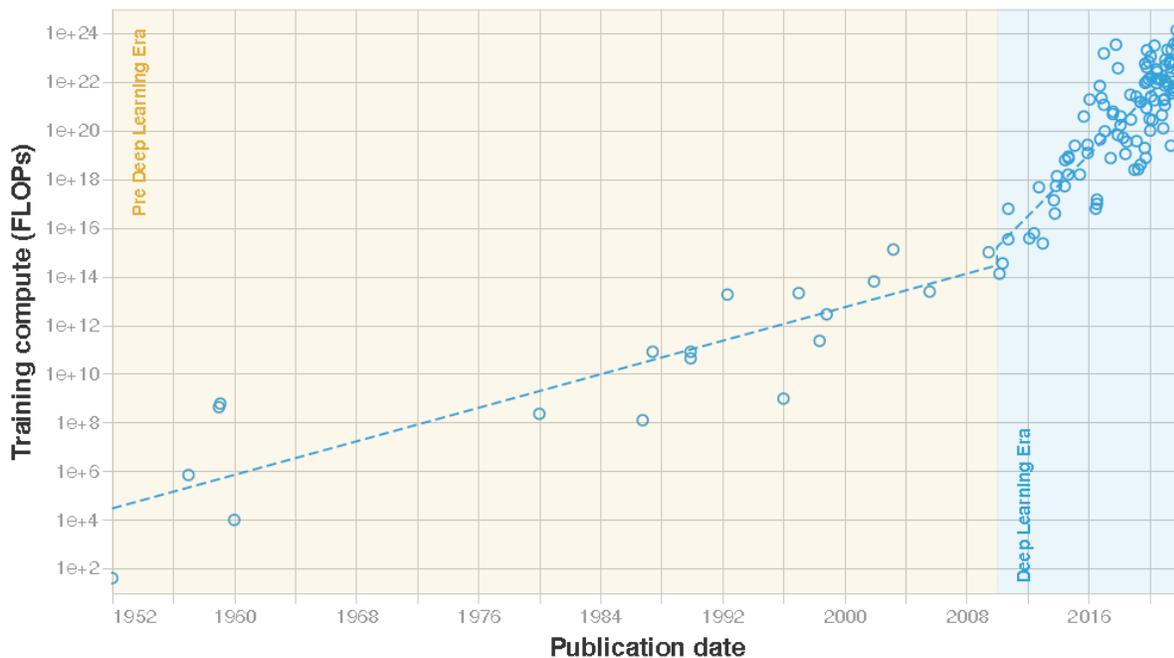
By [Charles Fadel](#) and [Wayne Holmes](#)

“We tend to overestimate the effect of technology in the short run, and underestimate the effect in the long run.” - Roy Amara, Former President, Institute for the Future

Our book [“Artificial Intelligence in Education”](#) is now 4 years old - an eternity for exponentially-driven technologies. It dealt primarily with the impact of early Machine Learning/Deep Learning (ML/DL) on Education.

But DL has progressed significantly faster than Moore’s Law, according to recent research¹: *“[Before DL], the amount of compute required to train ML systems doubled once every 17 to 29 months. Subsequently, the overall trend speeds up and doubles every 4 to 9 months. The trend in the Pre Deep Learning Era roughly matches Moore’s law, according to which transistor density doubles roughly every two years (Moore, 1965) – often simplified to computational performance doubling every two years.”* (note the log scale on the y-axis)

Training compute (FLOPs) of milestone Machine Learning systems over time
n = 121



¹ “Compute Trends Across Three Eras of Machine Learning” Jaime Sevilla et al, <https://doi.org/10.48550/arXiv.2202.05924>

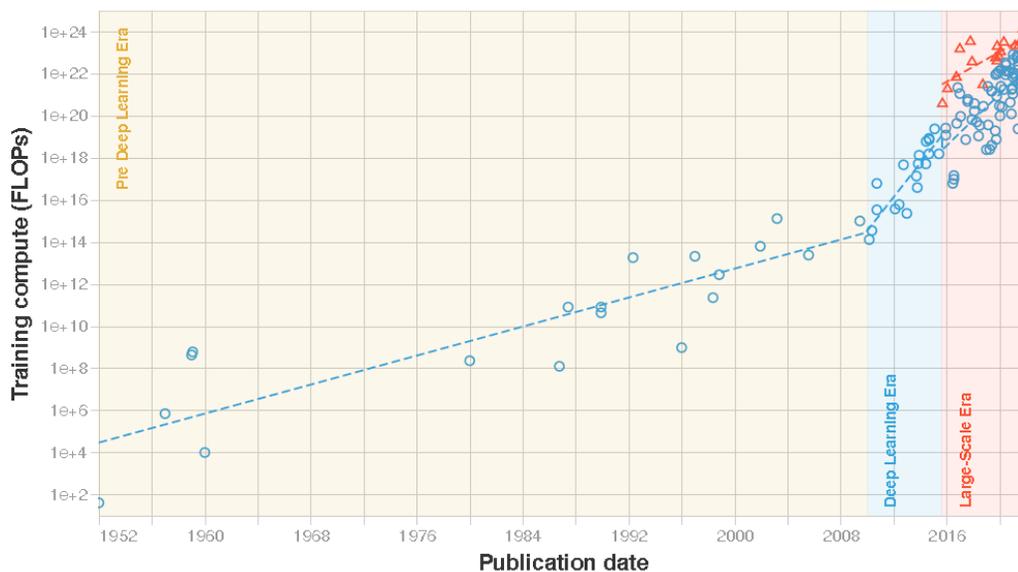
ML/DL systems have allowed ground-breaking innovations, a shining example of which is AlphaFold² from DeepMind, which can “accurately predict³ 3D models of protein structures and has the potential to accelerate research in every field of biology.”

In addition to natural-sounding speech⁴, they have mastered games like Meta’s Diplomacy⁵ and DeepMind’s Stratego⁶, which are a lot more complex than chess and Go, that, like Poker, require the *strategic ability of bluffing* →

	Chess	Poker	Go 19x19	Stratego
Turns in a game	~60	~15	~300	~1000
Starting configurations	1	10 ⁶	1	10 ⁶⁶
Game tree complexity	10 ¹²³	10 ¹⁷	10 ³⁶⁰	10 ⁵³⁵

However, the innovation that most captured the public’s imagination has been the emergence of “Foundation Models”⁷ (aka “Large Language Models” (LLM), “Transformers”), which have outpaced Moore’s Law by a factor of up to 20X^{8,9}, with their own distinct evolutionary curve.

Training compute (FLOPs) of milestone Machine Learning systems over time
n = 121



² <https://www.deepmind.com/research/highlighted-research/alphafold>

³ But it can't help in other stages, such as modeling how drugs and proteins would physically interact. https://www.theregister.com/2022/09/08/deepmind_alphafold_performance/

⁴ <https://www.deepmind.com/research/highlighted-research/wavenet>

⁵ Cicero. Meta AI (2022).

⁶ “Mastering Stratego, the Classica Game of Imperfect Information.” Perolat et al. (2022).

⁷ “A foundation model is a large artificial intelligence model trained on a vast quantity of unlabeled data at scale...resulting in a model that can be adapted to a wide range of downstream tasks.” https://en.wikipedia.org/wiki/Foundation_models

⁸ “AI Training Is Outpacing Moore’s Law” <https://spectrum.ieee.org/ai-training-mlperf#toggle-gdpr>

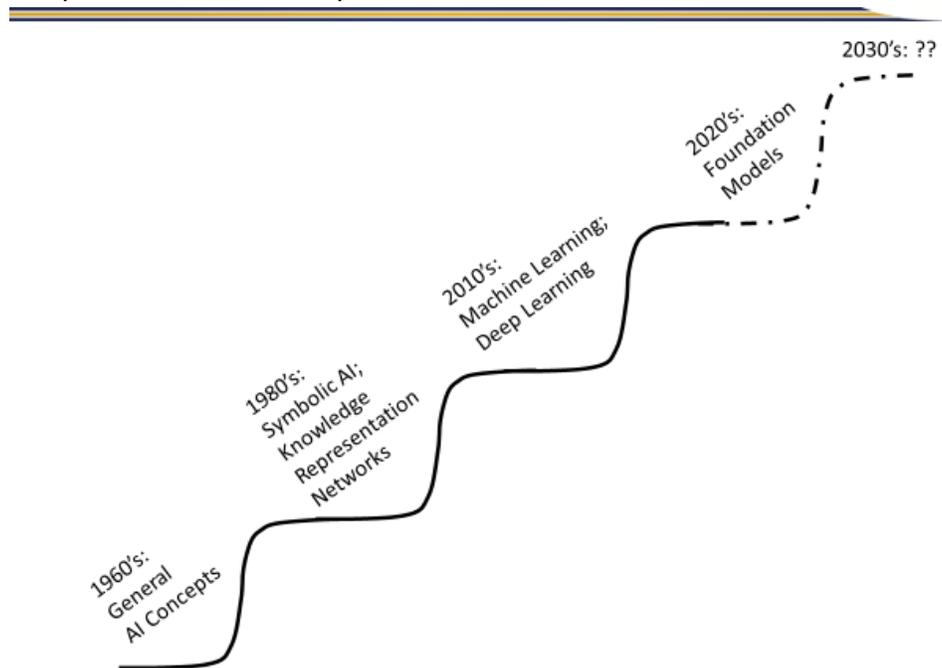
⁹ “Compute Trends across three eras of machine learning” <https://arxiv.org/pdf/2202.05924.pdf>

Training compute (FLOPs) of milestone Machine Learning systems over time

n = 102



FM's ushered a major step-up once they *passed a threshold* in the number of parameters. We are witnessing the sudden acceleration of a fourth generation AI, leading to another step in the “punctuated equilibria” of its development:



While GPT2 and related models were somewhat unimpressive, GPT3 has suddenly inflamed the imaginations with its apparent ability to “create¹⁰”:

- Writings: address the UK Parliament¹¹; Twitter threads; novels; summaries; poems; Q&A prompts for speakers; essays; naming; translating; lesson plans; computer code; commenting on social media; bias reduction; interview preparation questions; editing for AP style; explaining quantum physics concepts; etc.
- Advice: dating; anxiety; fitness; experts; content moderation; crime mysteries; etc.
- Designs: games; furniture; gift ideas; etc.

At this stage, FM’s benefits from the following advantages:

- Large size (trained on a massive corpus of text data). This allows the model to generate common opinions and fluent responses to a wide range of natural language processing tasks.
- Human-like responses: to generate responses that are highly relevant to the prompt and that exhibit a level of knowledge that appears to be similar to that of a human.
- Adaptability: to adapt to different situations and contexts, which makes it useful for a wide range of applications.
- Versatility: that can be used for a wide range of natural language processing tasks, such as generating high-quality text responses quickly and easily.

But FM’s have also been called a “bullshit generator that can still be amazingly useful¹²”. Their limitations stem from the fact that FMs are disembodied (having no proprioception or interoception), isolated (from other real-world information), and - so far - lack ethical conscience, due to:

- Dependence on data: the quality and accuracy of the model’s responses will depend on the quality and diversity of the data that it has been trained on.
- No understanding: no deep understanding of the world or the ability to reason like a human. As a result, the model cannot generate responses to complex or abstract questions, or to understand the context and implications of a given prompt.
- Bias: if the model has been trained on biased or unrepresentative data, or if the model’s algorithms are biased in some way.

There was a *significant* step-function in capabilities from GPT-2 to GPT-3, which became

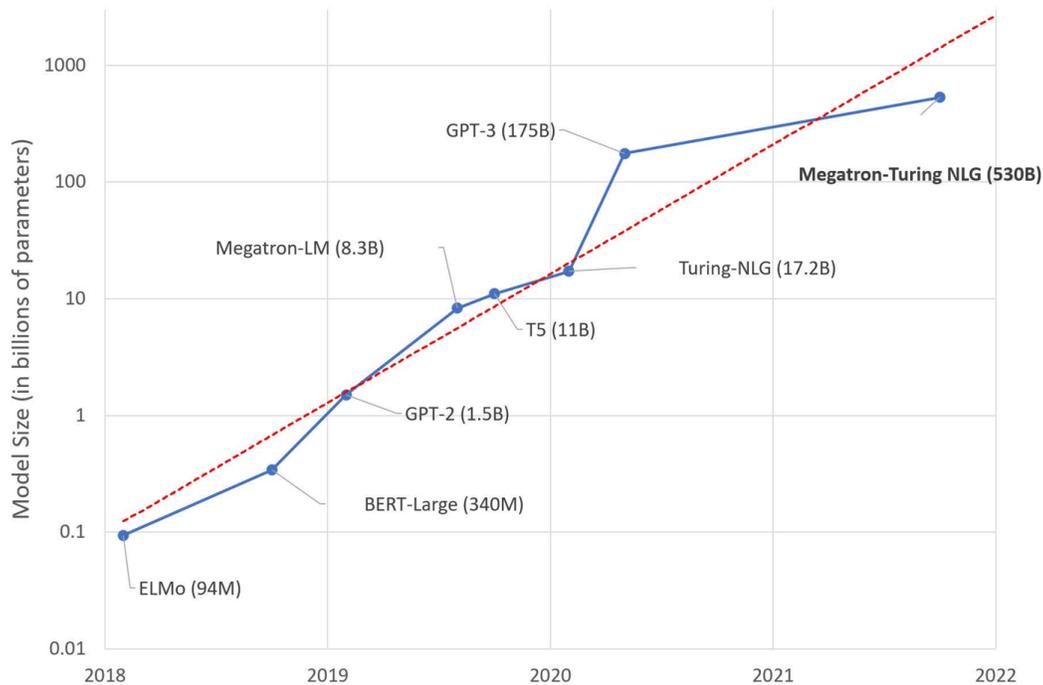
¹⁰ Create in response to human-written prompts, and only create through combining/correlating things that humans have created.

¹¹ <https://analyticsdrift.com/conservative-mp-delivers-an-ai-generated-speed-in-the-house-of-commons/>

¹²

<https://aisnakeoil.substack.com/p/chatgpt-is-a-bullshit-generator-but#:~:text=But%20it%20can%20still%20be%20amazingly%20useful,-Arvind%20Narayanan&text=The%20philosopher%20Harry%20Frankfurt%20defined.is%20the%20greatest%20bullshitter%20ever.>

particularly visible when the ChatGPT interface was offered to the public. One wonders what will happen with even larger models such as Megatron-Turing and GPT-4: since there was a threshold crossed between GPT2 and 3, will we continue to see similar progress, or will we see a *saturation effect* of FM’s capabilities? The next few months will be instructive.



Intelligence Augmentation in Education

Last century, biology transitioned from descriptive (observational) to analytical (DNA analysis). The same happened to a certain extent with the Arts, with the advent of many creation tools like Photoshop and visual effects in cinema, up to very recently with visual models like DALL·E¹³. As of this writing, Computational tools are increasingly tried in all domains, even poetry now¹⁴.

The rapid acceleration due to FM’s have left scholars and instructors anxiously asking for clarity about:

1. the true capabilities of the technology, at this stage, and ideally some extrapolations
2. its impact on the “humanistic social science” of Education
3. its affordances *as* EdTech in that context.

These are the questions that we will research; the intellectual significance lies in understanding and clarifying what is IA capable of, by exploring the boundaries between Reckoning (calculating

¹³ <https://openai.com/dall-e-2>

¹⁴ [GPT-3 Demo: Poetry](#) OpenAI (2022).

prediction) and Judging¹⁵ (practical wisdom) on one side, applied to bounded vs unbounded problems¹⁶ on the other side. The conversation thus will revolve around clarifying this matrix for Intelligence Augmentation →

	Reckoning	Judging
Bounded	AI	IA?
Unbounded	IA?	Human

The impact of a renewed analysis¹⁷ on Education is two-fold:

1. Clarifying the What: *“should students still write essays?”* is a typical question being asked by instructors.
2. Clarifying the How: *“how should students use FMs to write better essays?”* is a typical question that follows. And if so, how can AI as Education Technology help - how would the FMs be presented to the students, which ones, and in what format? What are the ethical safeguards to put in place?

This will require us to delve into epistemological questions: the nature of Understanding, Trust, and Transfer. The research questions to be asked are, as applied to the two IA quadrants in the matrix above:

1. What does it mean to Understand something?^{18, 19}
2. What does it mean to Trust something?²⁰
3. How can we promote Transfer?²¹

Finally, there are also two other issues to consider:

1. Embedded effects: once a technology is easy-to-use, it spawns an array of innovations: to use a computer analogy, ML/DL is operating-system-domain, while FM’s are application-domain. Another analogy here is the car: the user does not need to understand the engine in order to drive the car, but they do need to understand some mechanic (such as pressing one pedal makes it go faster and another makes it go slower), and more importantly they do need to know the Driving Code (the rules of driving safely). In other words, the user does need some critical awareness of what these

¹⁵ Private communication from Dr. Chris Dede, Harvard Graduate School of Education.

¹⁶ As discussed in engineering circles.

¹⁷ [State of the art and practice in AI in education - Holmes - 2022 - European Journal of Education - Wiley Online Library](#)

¹⁸ “The Debate Over Understanding in AI’s Large Language Models” M. Mitchell, DC. Krakauer <https://arxiv.org/abs/2210.13966>

¹⁹ “On the Measure of Intelligence” F. Chollet, Google <https://arxiv.org/pdf/1911.01547>

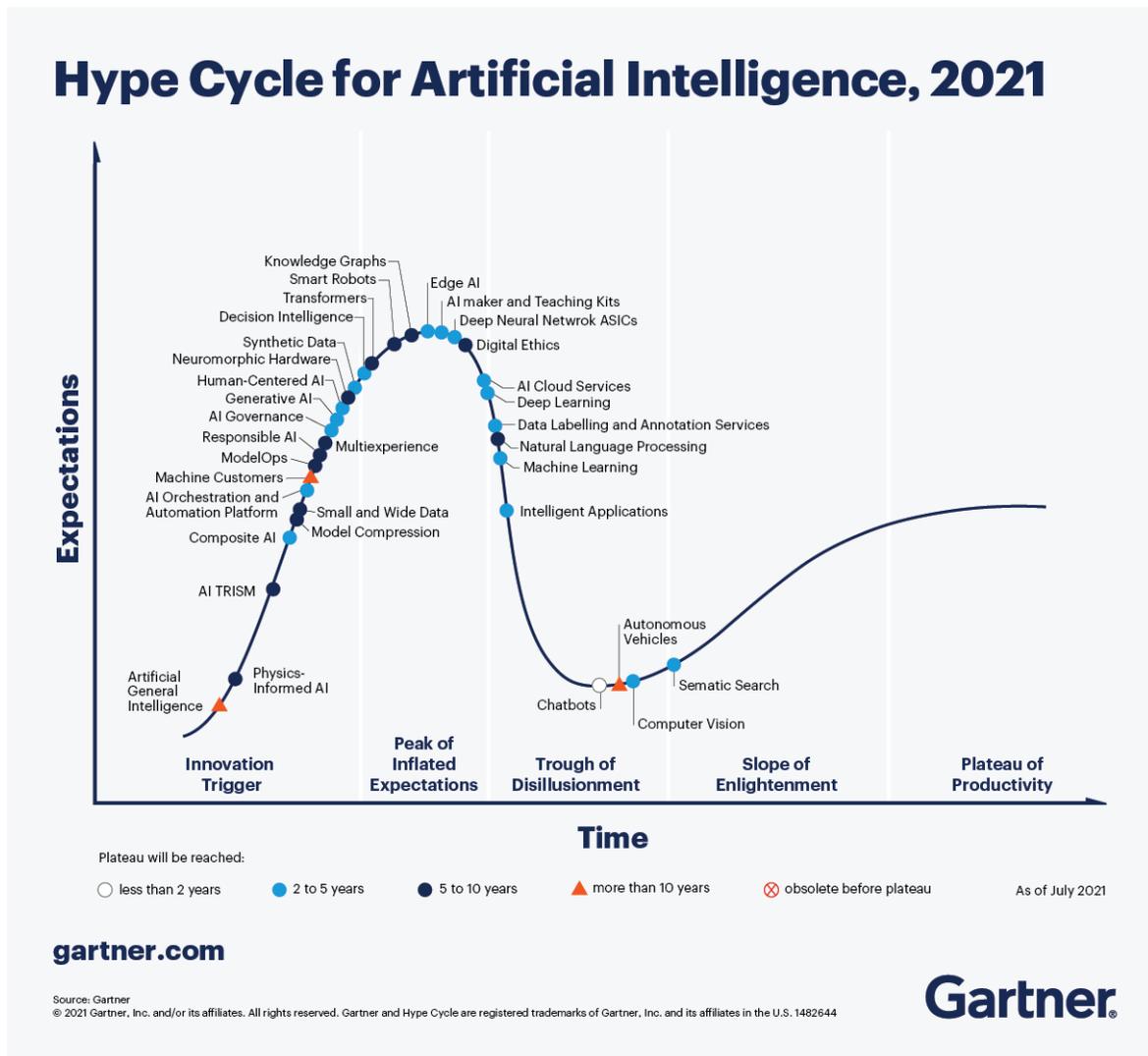
²⁰ “Three Models of Epistemic Dependence”, Benjamin McMyler, <https://doi.org/10.1093/acprof:oso/9780199794331.003.0004>

²¹ “Abstraction and Analogy-Making in Artificial Intelligence”, Melanie Mitchell <https://arxiv.org/pdf/2102.10717> and David Perkins’s “Making Learning Whole” (Jossey-Bass, 2010)

tools can and can't do, and about what they should be allowed to do and not allowed to do (e.g. predict recidivism).

2. Unintended consequences: when Short Messaging System (SMS) was created as part of the GSM cellular phone standard, no one forecast that a bit of unused guardband bandwidth between phone frequencies could be SO successful:
 - a. First, intrinsically, as a method of communication
 - b. Second, at modifying commitment via just-in-time communication
 - c. Third, to undermine Democracy itself, in its Twitter embodiment

So predicting the future is particularly dicey at this stage²², which is why we will not tackle it. Gartner had in 2021 considered FM's to be reaching "peak hype", only to be proven wrong...



²² "Prediction is very difficult, especially about the future" - Niels Bohr

One of the problems with ML/DL and FM's is that they have vastly outpaced the traditional scholarly literature publishing cycle, leaving us to extrapolate where the technology might get to. *It is clear that by the time our next book is written, the intense exponential growth of FM's will force us to reconsider the tenor of some of our questions, which should be viewed as not just acceptable, but desirable.*

In the words of a reviewer of our first book: *"AI is increasingly used in education and learning contexts. The collision of three areas – data, computation and education – is set to have far-reaching consequences, raising fundamental questions about the nature of education: what is taught and how it is taught."*
