

## AI's Recent History

### The Skinny on AI for Education, August 2023

#### Headlines

- [The Main Read: We must now futureproof our education systems for an AI world](#)
- **News:**
  - [Notes for the Future: Much more AI to come our way and worries about fake vs. reality](#)
  - [Generative AI news: More choice, changes to ChatGPT terms and conditions and added functionality](#)
- [Regulation and Ethics: Let's work together](#)
- [Policy and Debate: There is no consensus about the real risks of AI for humanity](#)
- [Let's put GenAI and ChatGPT in their place: The fascinating journey of AI: from automata to generative AI](#)

#### **The Main Read: We must now futureproof our education systems for an AI world**

Last week I was delighted with the reception received by an opinion piece I wrote for the Guardian newspaper. I have long believed that we need to rethink our education systems to ensure that we better prepare people for a world that will inevitably be full of AI of different types, performing a range of functions. This means that we must act now to help everyone understand more about AI: what it is, what it can do, what it can't do, and how importantly different it is from human intelligence. And we must develop a future proof education system. This will be a system that pays great attention to helping people develop the higher order sophisticated thinking skills that we humans are capable of achieving and AI is not. The additional advantage of doing this is that developing these thinking skills, like metacognition and self regulation, also helps us become better at learning about the world and demonstrating our knowledge of the subject areas that we hold so dear through our existing examination systems.

The real beauty of the innovations in AI that have been bubbling away for the past few years is that we can now also assess the development of these higher order thinking skills - we can now use AI to measure what was previously considered unmeasurable and that ought to be a powerful prospect for re-thinking assessment.

In the past such views as those that I expressed in this article have often been received with scepticism and in some cases complete (and energetic) rejection. Even as recently as this spring, when I gave evidence to the UK House of Commons Science and Technology Select Committee for their inquiry into the Governance of AI, my views precipitated some significant disagreement. However, I feel that the objections to such views are starting to recede and that can only be a good thing for possible reforms.

It is already becoming very clear that the world of work has and continues to change and the implications for education and training are profound. For example, the 2023 World Economic Forum Future Jobs report stated that "Employers estimate that 44% of workers' skills will be disrupted in the next five years." Self-efficacy skills, curiosity, lifelong learning; resilience, flexibility and agility; and motivation and self-awareness are key areas of demand and training need. The report also highlights that "6 in 10 workers will require training before 2027, but only half of workers are seen to have access to adequate training opportunities today." Surely this is a call to action that we must start to respect.

One only has to look at Estonia and Singapore for examples of countries that are 'ahead of the game' when it comes to AI and enabling a more holistic learner focussed education system.

#### **News: Notes from the future and developments in generative AI**

It is always tricky when writing about AI to get the right balance between enthusiastic optimism about a technology that could bring such great benefits to education and the cautious measured thinking and action that is necessary when it comes to a technology that is not well understood, even at times by those who are building it! In this section I discuss a little about specific types of technology in an attempt to set the scene. However, whilst I encourage everyone to learn about AI and to try things out, I would not want anyone reading 'The Skinny' to feel pressured into using AI or to feel inadequate, because they are not up to speed with all the latest tools.

**Notes for the future: much more AI to come our way and worries about fake vs. reality:** One of my favourite newsletters is 'The Batch', which ran a fascinating piece last week about where money is being invested in AI companies. Whilst the data discussed only relates to the top 100 AI start-ups as judged by CB insights it is still fascinating and reveals that there has been another \$22 billion in funding devoted to AI and in particular generative applications - so expect many more tools and options to appear on the horizon over the coming months and improvements to those that already exist. Expect more AI assistants and tools to support search and producing computer program code.

On a worrying note, I noticed the increased investment in companies specialising in hyper-real deep fake videos. On 8th June this year Gillian Tett in the Financial Times reported the use of a deep fake video of a explosion near the Pentagon.

**Why this matters for education:** I worry that deep fakes of apparently knowledgeable and respected people will appear presenting plausible but incorrect information for students. It increases my belief that we must help our students become better at making critical judgments about what they should believe, including understanding how to seek and find corroborating evidence to help them decide between truth and fiction - this is also reflected in the Guardian Article I discuss at the top of 'The Main Read' section of this newsletter.

**Generative AI news: more choice, changes to ChatGPT terms and conditions and added functionality:** Some free tools to explore include Claude and Pi. Claude is produced by Anthropic, who specialise in AI products that emphasise safety. Bard is also improving, Elon Musk launched xAI, to challenge the dominance of OpenAI, and Meta released an open source Large Language Model called Llama 2, for public use. This means that anyone with the inclination and know-how can build a ChatGPT like chatbot. Another reason to expect ever more chatbots to come our way. Being open source makes such technology available to many people, but it also adds a risk because not everyone will behave responsibly with this powerful technology.

It is useful to try a range of tools to see which you find easiest and most suitable for the tasks you need to do. It is also nice to have a choice of options and that situation is only set to increase, as our 'Notes for the Future' section highlights. But of course there will be increasing costs associated with these. For example on 18th July, Microsoft announced that it will charge \$30 a month for the generative artificial intelligence features in its software.

**Changes to ChatGPT terms and conditions:** A change to the ChatGPT terms and conditions means that you can now turn off your history and prevent your conversations being used to train OpenAI's models. OpenAI's ChatGPT now allows users to disable the chat history feature, meaning the chatbot will not save any previous conversations once this setting is turned off. Furthermore, OpenAI explicitly stated that it will not use these conversations to train its AI models. While new conversations will still be stored for up to 30 days for monitoring abuse, they will be permanently deleted after that period.

**Why are the changes to ChatGPT relevant to education?** The change is not a "cure all" for ethical concerns associated with these technologies, but it is a step in the right direction, although it does mean that once history is turned off you can't revisit past conversations. For educators who use AI-powered tools like ChatGPT in educational settings, data privacy is a critical concern. By turning off chat history, educators can ensure that any sensitive information or student interactions are not stored or used for model training, safeguarding both their students' privacy and their institution's data. Students may feel more comfortable expressing their thoughts and seeking help without the fear of their interactions being recorded and analysed. By not utilising chat history for model training, OpenAI also reduces the

risk of perpetuating biased or inaccurate information in the AI's responses. This benefits educators by ensuring that the AI remains a reliable and unbiased resource for their students. ChatGPT will still retain new conversations for 30 days for abuse monitoring. After 30 days, the conversations will be permanently deleted.

**Key takeaway:** *new technology innovations always change and evolve quickly so expect more changes. See this as a positive step, but still be cautious about what you enter in these tools - do you have a right to the material you are entering? Are you compromising the IP of the text that you enter? And remember, this change does not make ChatGPT more accurate - it will still make things up!*

**Code Interpreter:** OpenAI also recently introduced a new feature called "Code Interpreter" but only for paying 'ChatGPT Plus' subscribers. This feature enables ChatGPT to write and execute computer code to find the answer to a question you pose, and this means that it can generate charts based on your user-uploaded data – it should also increase the accuracy of the tool, but bear in mind that there will still be mistakes. Code interpreter enables the production of basic data visualisations and provides some data analysis functions too. EdTech Insiders' newsletter has a nice article about this from an educational perspective.

### **Regulation and Ethics: Let's work together**

**The EU AI Act:** The EU AI act continues its progress through the EU Parliament, Commission, and Council of member states — and we can only wait and see how this transpires and what changes are still to come. The action on regulation over the past couple of weeks has focussed more on the US than the EU. US lawmakers are conducting a comprehensive review of AI, aiming to strike a balance between its benefits and potential risks. And, as reported in the previous issue of The Skinny, US Senate majority leader Chuck Schumer has proposed expert briefings and forums for key senate committees to determine which aspects of AI require regulation.

A few days ago, the Federal Trade Commission launched a comprehensive probe into OpenAI, the maker of ChatGPT. The investigation will assess potential harm caused by the chatbot's creation of false information about users and it will examine OpenAI's privacy and data security practices. Additionally, the FTC has requested OpenAI to disclose the training data for its large language models, a request that OpenAI has thus far refused. This is a space to watch for indications about the power of Silicon Valley.

On 23rd July an interesting Op-Ed was published by Antony Blinken and Gina Raimondo, US Secretary of State and US Secretary of Commerce in which they establish that the US will “partner with countries around the world, as well as the private sector and civil society, to advance a key goal of the commitments: creating AI systems that make people's lives better.” and that “to shape the future of AI, we must act quickly. We must also act collectively. No country or company can shape the future of AI alone. The US has taken an important step — but only with the combined focus, ingenuity and co-operation of the international community will we be able to fully and safely harness the potential of AI.”

### **Policy and Debate: There is no consensus about the real risks of AI for humanity**

**An open letter to counter AI doom:** More than 1,300 experts call AI a force for good. Organised by BCS, the Chartered Institute for IT, the open letter was intended to 'counter "AI doom"'. The BCS supports the need for rules around AI, but takes a more positive view than those who signed the letter claiming AI was a threat to humanity.

I can certainly sympathise with their point of view and believe that the biggest risk is that posed by ignorance, which may lead us to believe that AI is far more intelligent than it really is, thus causing us to allocate it responsibility for actions that should continue to be done by humans. Having lived through the AI Winter at the end of the last millennium and the start of this one, I do not believe that Artificial General Intelligence (AGI) is going to be with us any time soon, because each time we think we think we are close, we realise that once again we have misjudged exactly how complex human intelligence really is and that actually we are nowhere close. For example, the extent to which any AI is capable

of behaviours that might be considered metacognition is so rudimentary that we would not even credit it with that name if we observed such behaviours amongst humans.

### **Let's put generative AI and ChatGPT in their place: The fascinating journey of AI - from automata to generative AI**

The technology behind the generative AI tools that are flooding the market is not new; it's time to find out more about the AI timeline!

***The fascinating journey of AI: from automata to generative AI:*** Long before the 1920s, people were captivated by automata and humanoid figures that seemed almost magical. Little did they know that these early manifestations would lay the groundwork for the exciting world of AI. In 1920, Karel Čapek introduced the word "Robot" to our language, giving a name to the concept of intelligent machines that would become central to AI's development.

Fast forward to the 1940s, when Norbert Wiener embarked on the birth of cybernetics, a field that explored the study of control and communication in both animals and machines, marking a significant step towards understanding AI. In the 1950s, AI took a giant leap with Alan Turing proposing the famous Turing Test, asking the question, "Can Machines Think?" This test evaluated a machine's ability to exhibit intelligent behaviour similar to humans, sparking the birth of modern AI.

A pivotal moment came in 1956 with the Dartmouth College meeting, where a group of 12 visionaries officially established AI as a field of research, setting the stage for its remarkable growth. In 1957, Frank Rosenblatt's perceptron, the first neural network, emerged, unlocking new possibilities for AI and laying the foundation for future advancements.

Throughout the 1960s, chatbots began to emerge, and the first one, called Eliza, fascinated users with its ability to engage in human-like conversations. The 1970s witnessed the rise of Expert Systems, which relied on rule-based methods and statistics. These systems showcased impressive problem-solving abilities in specialised domains. The pinnacle of these expert systems was a machine called Deep Blue, which achieved a historic victory on May 11, 1997, by beating the then chess world master Gary Kasparov at chess. This event marked a significant milestone in the world of AI, and for a moment, it seemed like AI had finally cracked human intelligence.

However, the AI Winter that followed brought a reality check when researchers realised that AI hadn't cracked intelligence entirely. It was apparent that achieving human-like vision, for example, was far more challenging than mastering games like chess. But, the early 2000s brought a renewed interest in AI with a focus on machine learning using neural networks and statistical methods. In 2009, the introduction of Convolutional Neural Networks (CNN) paved the way for superior performance in tasks involving images, speech, or audio signals. That same year, Fei-Fei Li created Imagenet, a vast dataset that mirrored the real world, providing valuable training data for AI models.

The year 2011 was momentous as Apple integrated Siri into the iPhone, revolutionising how we interacted with our devices. Additionally, IBM's Watson wowed the world by defeating human champions on Jeopardy, showcasing AI's problem-solving abilities. In 2012, Alex Krizhevsky created AlexNet, the first CNN based on Fei Fei's model, pushing the boundaries of image recognition and setting new standards for AI performance.

The year 2014 saw a chatbot named Eugene pass the Turing Test, blurring the lines between human and machine communication, while Amazon launched Alexa, an AI-powered virtual assistant that became a part of many households. Another watershed moment for AI came in 2017 when Google's AlphaGo made headlines by defeating the world champion at the ancient board game Go, solidifying the rise of Deep Neural Networks as a prominent trend in AI.

In 2018, the first large language models (LLMs) were created, showcasing impressive language understanding capabilities. And in 2022, the LLM, ChatGPT was released, representing a breakthrough in AI, not because the technology was new, but because it was the first time that sophisticated AI had been freely available to billions of people through an easy to use interface.

Commercial LLMs began to proliferate, ushering in a new era for AI particularly in education. Where this journey ends will depend upon the decisions we humans make now and to make good decisions we all need to understand enough about AI – so let's get learning!