AI & Youth Work

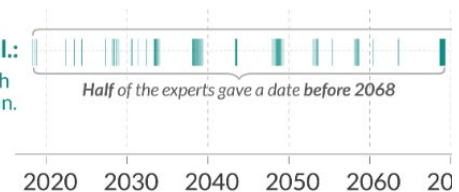## A briefing paper to support Youth Scotland members

Contents

Executive Summary

This briefing paper's purpose is to present the findings from research conducted into Artificial Intelligence (AI) and its potential applications for youth work organisations.

The main finding of this briefing is that while AI is increasingly available in personal and professional situations, it is currently only one such tool and, like all tools, only useful in appropriate application and use cases. It is, however, likely to change our lives in the next decade. Due to technical, ethical and sustainability limitations, it is not yet the panacea many have claimed but it is a real and present impact on daily life. It is the emergent and disruptive technology of the moment, already affecting the workplace and requiring inclusion in any organisation's strategic and operational planning. While predictions are difficult, 50% of AI experts think that 99% of current tasks humans are paid to do could be accomplished by AI systems in the next 40 years.

Fig1: Expert view on AI systems timeline. Source: Our World in Data, 2024.

Due to the nature of the topic, jargon and the use of acronyms is regrettably unavoidable. There is a glossary of key terms and product names in Appendix A.

Disclosure

The cover image of this paper was generated by DALL-E, a ChatGPT image generative AI, based on a Youth Scotland photo. The document was shared with ChatGPT to suggest anything a generative AI might consider to be missing (suggestions were the bias reduction and expert networks). ChatGPT also provided an expanded glossary, based on provided definitions. All three documents were written by a human.

Scope

The briefing paper is intended to offer background research and share information, giving an overview of AI, as relevant to the youth sector. It also shares how Youth Scotland is approaching AI and what guidelines we are using to do so. It is not intended as a *fait accompli* but as a basis with which Youth Scotland – and hopefully our members - might engage appropriately with in an incredibly fast moving field.

Our intention is to release this document to membership and engage in a listening and consultation period to better understand member groups and youth workers who may have concerns or are interested in understanding more from a community-based youth work perspective. What is good for Youth Scotland, should be good for our members.

Background and Context

In spring of 2023, the terms 'AI' and 'ChatGPT' exploded into the public consciousness, claiming nothing less than new technological advances put society on the verge of another industrial revolution. The estimated value of OpenAI, ChatGPT's parent company, was at $86bn in October 2023. There were daily stories online and in broadcast media, as an arms race unfolded between the big tech companies to develop or include AI in their products. ChatGPT, DALL-E, Gemini, Bard, Copilot and more were entering the public consciousness as tools that could complete tasks quicker and better than we could, promising better efficiency, quality and results.

While many of the Large Language Models (LLMs) and AI powered tools have been available for some time, two key changes in 2024 have driven this paper: the freely-available access to these tools and corresponding media coverage puts AI on the table for the youth work sector like all others; secondly, the recent rush of tech companies to build AI into their existing apps and services – such as Copilot with Microsoft Office or other GPTs in search engines – makes their consideration unavoidable and essential, particularly as they are not fully understood nor well-tested.

Fig2: Google Trends data for relevant Google searches – August 2021-2024. Source: Google Trends, 2024.

Start With Why

Why Should We Care?

Benefits could be significant for the youth sector

Capacity, volunteer time, funding and reporting are all well-known pain points for community-based youth work – what if AI could help write a funding application and even assist with the reporting back to the funder? Youth workers could use AI to help prepare session plans, write governance policies, create content for their social media, interpret reports and other documents and so on. Where this isn't wholly possible for an AI, it may at least be able to reduce the time taken.

Young people care

During the Scottish AI Summit 2024, a youth panel provided fascinating insight into what young people in Scotland think about AI. The panel was hosted by YoungScot and Dr Mhairi Aitken, Ethics Fellow in the Public Policy Programme at the Alan Turing Institute, and shared some impactful statistics from young people:

- 48% of young people use AI
- 63% of those use it to get answers to questions
- 33% believe AI can get the right answer to things
- 17% use it to chat to someone they know isn't real

Some of the responses could raise concern, concerns many of those young people share:

- 54% worry about what happens to their information
- 53% worry about how AI uses their data
- 51% worry about privacy
- 47% worry about how much AI knows about them

Regardless of our wider feelings on AI and its adoption, these responses show that young people are regularly using AI in common ways, despite their concerns. As per the Nature and Purpose of Youth Work, 'Youth workers form developmental relationships with young people that help young people to explore their values and beliefs; address issues that affect them […]'

So, to remain relevant and able to maintain a trusted adult status, at a base minimum we should understand the technology, how it affects young people and how they are engaging with it, even if we do not wholesale adopt it in the delivery or running of youth work.

Why Do Companies Say 'Artificial Intelligence'?

AI refers to the simulation of human intelligence by machines, particularly computer systems. It involves the ability of machines to perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making and language translation.

While AI has been in development for many years, the recent media buzz has mostly focused on Large Language Models (LLMs) which are programmes that have been trained on huge amounts of written data to interact and respond to prompts in human-like ways. In this case, LLMs use Machine Learning (ML) to use algorithms and statistical models so they learn from and make decisions based on data. Unlike traditional programming, where explicit instructions are given, ML allows the system to improve its performance on tasks over time as it is exposed to more data. Our very act of using it is teaching it to improve.

LLMs have captured the public consciousness as they are Generative AI (GenAI), which means they can take a natural language prompt and create new text that mimics a human response. Equivalents now exist that can do the same with illustrations, photos, music and even video. As a simple example of text generation:
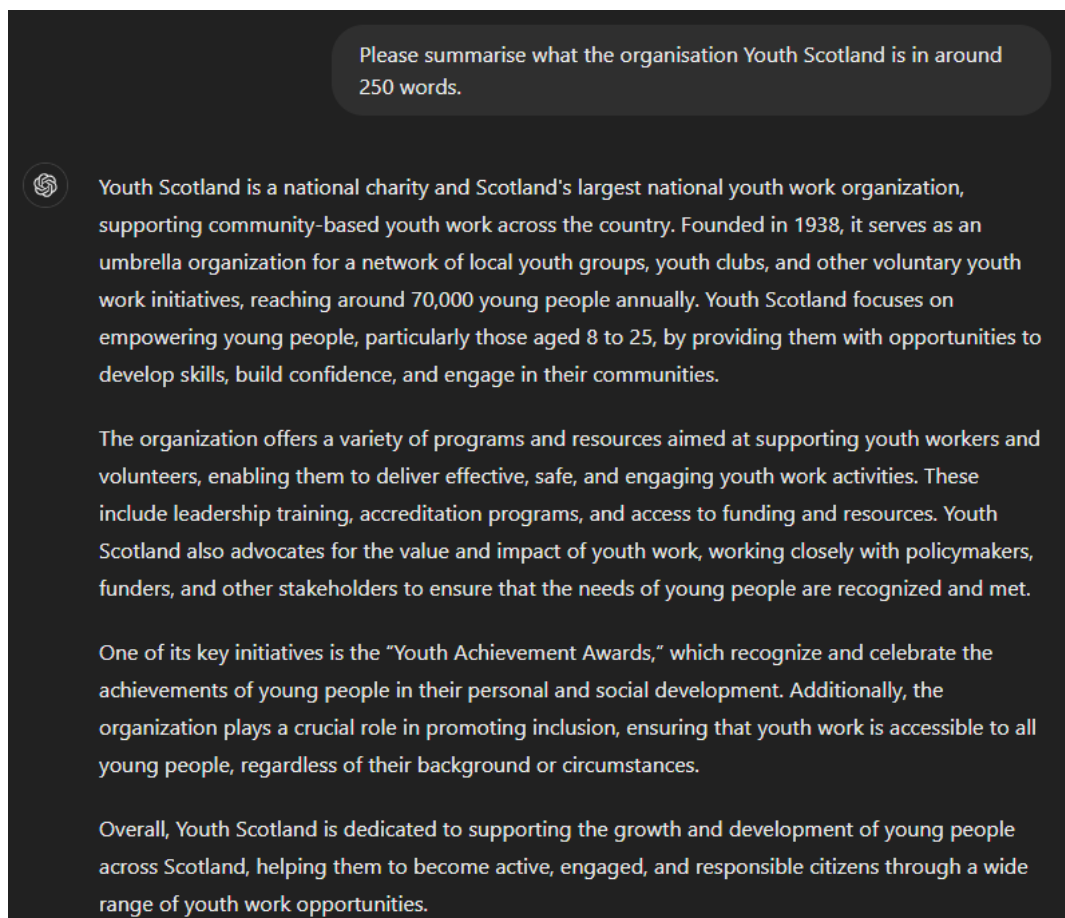

Fig3: ChatGPT response to a basic prompt. Source: ChatGPT 4o, 2024.

While there are a number of basic errors in the above demonstration – American spelling on a clearly UK-based prompt/subject, inaccurate facts like dates and volumes of young people supported – the output is remarkably human. ChatGPT 4o can use the internet to bolster its database, so while much of the response is close to what we have on our website, this answer is almost wholly new text, generated by an AI to sound as human as possible.

Generative AI like this can be combined with Cognitive Computing to interact with people in real time. For example, the Sainsbury's customer service chat on their website (and X/Twitter direct messages) is entirely AI unless an escalation occurs. Most interactions are conducted end-to-end without a human employee involved.

Is AI actually intelligent? No. Not by any conventional definition - AI is not self-aware nor does it understand the tasks it is fulfilling when prompted. AI makes for great marketing – certainly better than 'Cognitive Computing' or 'Large Language Models' – but these are not intelligences at work. In an oversimplification, experts describe AI's as being prediction and guessing models – models that guess based on vast quantities of similar data ensuring frequently accurate guesses. Perhaps more worryingly, AI developers do not yet seem to understand how their creations manage even this. In a recent scientific report commissioned by the UK government, it was stated that AI developers 'understand little about how their systems operate' and that scientific knowledge is 'very limited.'

Why Would Youth Charities Consider Using AI?

Charities are often operating with extremely limited capacity and/or limited access to skillsets out with their core function e.g. IT skills, data science or creative skills. Some potential applications could be:

- Operational efficiency: streamlining labour-intensive processes by automating repetitive tasks, streamlining data management/analysis, and improving communication/writing.
- Reporting: creating structure, analysing data, summarising feedback, sentiment analysis and infographic generation.
- Creative uses: image generation, advert generation, poster generation.
- Brain storming: helping smaller groups have bigger ideas or to think outside their own native experience and skillsets.

Risks and Challenges

- AI tools are often referred to as a 'black box' meaning that providers do not explain what happens to the data that is loaded into them. This presents a risk and we should not load anything that identifies another person, any sensitive or privileged information.

- Bias is a well-established concern of AI. An AI is only as knowledgeable and as neutral as the data it has been trained on. We should assume that most AI will have implicit bias of North American and western/white culture for many older models and, particularly now many have internet access, the various sources of human bias we see on the internet.
- Accuracy is a known issue with AI. AI's are known to 'hallucinate' which is colourful language meaning that they simply lie and make information up. Providers often refer to this as 'temperature' and responses should be requested with a low temperature when creating content or providing factual outputs. High temperature responses tend to be quite wild in their accuracy and interpretation.
- Connectedness – many if not most AI's now are connected to the internet in real time, meaning that they could potentially cross-reference data uploaded; think twice about how easy anonymised data could be completed by publicly available (online) information.
- Data privacy and security are key issues. As most AI's are a black box solution, or are potentially being trained on data that was not shared for this purpose, there are major concerns around GDPR and wider data protection.
- Trust: issues around the 'black box' nature of some AI systems, their willingness to 'hallucinate' and give misleading or incorrect responses and their general inhuman nature can lead to a lack of trust.
- Limited accessibility, either to internet access or the hardware to use it, presents a risk of small youth groups having a barrier to using AI.
- Sustainability issues are not just tied to the ethical consumption/production of silicon or hardware, but the extreme amounts of energy that AI systems and the related data farms take to power them. OpenAI has gone on record saying that the full capability of AI cannot be realised until we make a similar leap in energy production, such as Nuclear Fusion. The concern is here now and only going to get worse.
- Relatability/poor quality of generated content. Generated content frequently comes across as fake or bland – there is already a backlash against AI content and art.
- Abuse: 'deepfake' images, fake news and other abuse of powerful creative tools. If there is a way to break the law using digital tools (or otherwise) AI can probably be used to assist. Many AI's are written with safety protocols in place to prevent this (using a real person's likeness in a fake image for example) but users have gotten clever in circumventing this.
- Deskilled workforce – in a sector that values lifelong learning and the development of the young workforce, we must be careful not to sacrifice resourcefulness, core skills and problem-solving. We lose what we don't practice, after all.

- Long term impact on sector jobs is a valid concern. AI is already approaching human levels of competency in many areas, surpassing them in some. This is an issue for the working world in general. As with the move to more sustainable jobs, just transition of the workforce should be a consideration.

Test scores of AI systems on various capabilities relative to human performance

Within each domain, the initial performance of the AI is set to –100. Human performance is used as a baseline, set to zero. When the AI's performance crosses the zero line, it scored more points than humans.
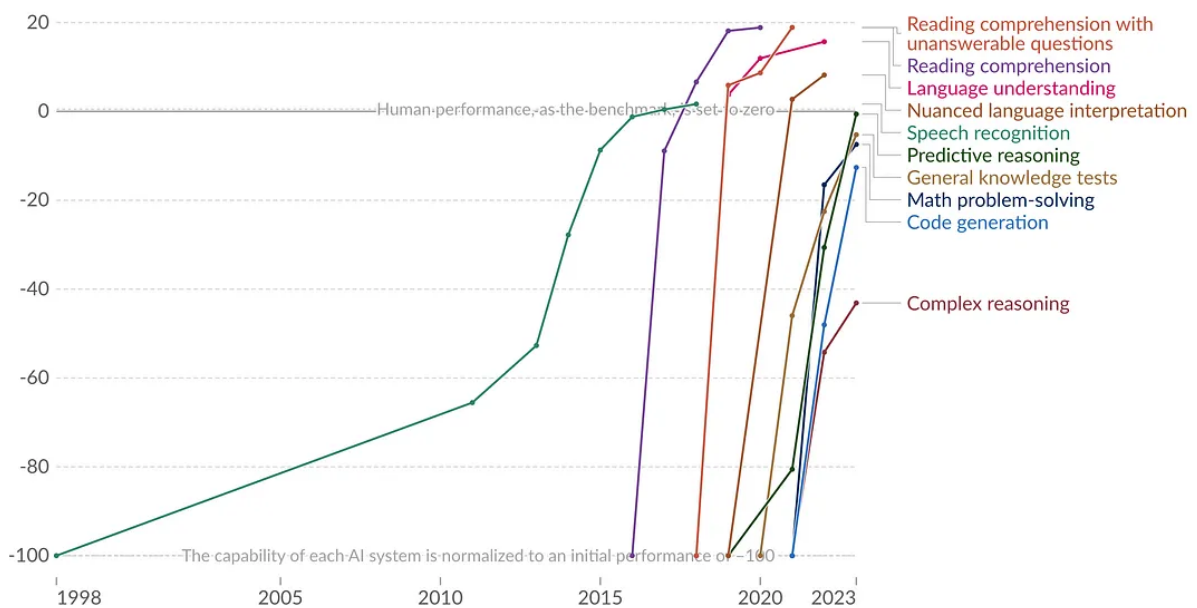
Data source: Kiela et al. (2023)
OurWorldInData.org/artificial-intelligence | CC BY
Note: For each capability, the first year always shows a baseline of –100, even if better performance was recorded later that year.

Fig4: AI capability testing 1998 – 2023. Source: Our World in Data 2023.

Risk Mitigation Strategies

Knowing the risks is part of the battle in an information-heavy subject. While the tools and their rapid development is increasingly hard to keep pace with, some basic strategies could help groups leverage the benefits more safely. Youth Scotland is taking the following steps:

Policy and guideline protection: we are implementing a set of guidelines and a robust policy to protect staff, ensure compliance with GDPR and Data Protection.

Bias Reduction: by understanding what biases may exist, we can aim to use diverse datasets, create processes for checks and balances that check outputs for obvious bias and establish clear policies for inclusive and equitable service delivery.

Ethical Frameworks: we can make use of global ethical frameworks (see Ethical Use below) advocate for transparent/ethical AI design with policymakers, implement our own regular process audits, and establish a committee/working group for review.

Training and Education: share learning and offer training to staff and member organisations about AI risks and responsible use.

How – Implementation Guidelines

Balanced, proportional, ethical.

To guide us as we explore this new area, Youth Scotland is working to a simple guideline that any usage or AI integration is balanced, proportional and most importantly, ethical.

Balanced – while AI is the new and exciting tool, it should be balanced with tried and tested methods, knowledge and other tools. Use the appropriate tool for the job at hand and consider the impact and effect of introducing AI to an established process.

Proportional – AI is not the answer to all challenges and should be considered as whether AI would simplify or complicate the solution. The tool should relate to the task or 'don't use a sledgehammer to crack a nut'. Asking an AI to summarise a one-page report is unlikely to be proportional use of the tool compared to asking it to do the same with a 40-page report and appendices. We should also only use it for the purpose we intend, helping us stay in line with GDPR.

Ethical – how you make use of AI and for what purposes is as important as how you conduct yourself in all your work for Youth Scotland. In other words, how you do anything is how you do everything. See below for further thoughts from UNESCO on what ethical AI looks like. UNESCO has released a 10-point framework for worldwide adoption.

Beyond these guiding principles, we might consider the following for any implementation of AI:

- Needs assessment: conduct assessments to identify where AI can add value without disrupting core delivery.
- Pilot projects: start with pilot projects before scaling up, adjusting based on successful outcomes and learning.
- Policy development: collaborate with stakeholders to form comprehensive AI policies.

Ethical Use

AI usage should reflect the ethos and aims of the youth work sector as a whole. A useful – and fairly comprehensive – guide to what ethical AI looks like can be seen in the [UNESCO Global AI Ethics and Governance Observatory](#):

"The aim of the Global AI Ethics and Governance Observatory is to provide a global resource for policymakers, regulators, academics, the private sector and civil society to find solutions to the most pressing challenges posed by Artificial Intelligence."

A human rights approach to AI

Ten core principles lay out a human-rights centred approach to the Ethics of AI.

1. Proportionality and Do No Harm

The use of AI systems must not go beyond what is necessary to achieve a legitimate aim. Risk assessment should be used to prevent harms which may result from such uses.

2. Safety and Security

Unwanted harms (safety risks) as well as vulnerabilities to attack (security risks) should be avoided and addressed by AI actors.

3. Right to Privacy and Data Protection

Privacy must be protected and promoted throughout the AI lifecycle. Adequate data protection frameworks should also be established.

4. Multi-stakeholder and Adaptive Governance & Collaboration

International law & national sovereignty must be respected in the use of data. Additionally, participation of diverse stakeholders is necessary for inclusive approaches to AI governance.

5. Responsibility and Accountability

AI systems should be auditable and traceable. There should be oversight, impact assessment, audit and due diligence mechanisms in place to avoid conflicts with human rights norms and threats to environmental wellbeing.

6. Transparency and Explainability

The ethical deployment of AI systems depends on their transparency & explainability (T&E). The level of T&E should be appropriate to the context, as there may be tensions between T&E and other principles such as privacy, safety and security.

7. Human Oversight and Determination

Member States should ensure that AI systems do not displace ultimate human responsibility and accountability.

8. Sustainability

AI technologies should be assessed against their impacts on 'sustainability', understood as a set of constantly evolving goals including those set out in the UN's Sustainable Development Goals.

9. Awareness & Literacy

Public understanding of AI and data should be promoted through open & accessible education, civic engagement, digital skills & AI ethics training, media & information literacy.

10. Fairness and Non-Discrimination

AI actors should promote social justice, fairness, and non-discrimination while taking an inclusive approach to ensure AI's benefits are accessible to all.

Support for Youth Groups

Youth Scotland exists for its members and the adoption of AI is no different. We know our wider team will likely have many suggestions for opportunities to support members, some of which include:

- Listening and consultation – check in regularly with our members to understand current and emerging areas of concern they are dealing with in community-based youth work.
- Share learning – as Youth Scotland and members identify best practice and learning from exploration and testing, we can develop resources that explain AI in practical and accessible ways and signpost to best practice advice and materials that are available now.
- Funding assistance – help small groups identify funding for tech upgrades and/or training opportunities to upskill.
- Expert networks – connect groups with experts to assist in implementing AI responsibly, upskilling opportunities, funded training, pro-bono consultancy.

Conclusion

As the saying goes, Pandora's Box is open. Whether we like it or not, AI is here and it's not going away. The technology has developed faster than many governments can comprehend, never mind legislate for. There has been too much marketing hype, too

much investment and preliminary stock valuations combined with a literal arms race at a national security level.

As a cautionary note, we would recommend not dismissing where AI is currently. While we do not realistically know where developments are on a sigmoid curve, we would caution against dismissing AI content that is unrealistic or unconvincing as the sum total of what it can/will be – where we are now is the 'stupidest' that AI will ever be. The point of these AI's is that they learn from large datasets and from interactions with users – these two sources are increasing exponentially.

As stated at the outset, the purpose of this paper was to provide enough summation of background research and share information, giving an overview of AI, as relevant to the youth sector. It also shares how Youth Scotland is approaching AI and what guidelines we are using to do so.

We hope you find this paper useful and we would encourage members to share their own experiences, concerns and learning with us. We aim to be transparent in our own journey and share what works and what doesn't with the wider membership – please feel free to contribute to that process by attending one of our consultations or emailing comms@youthscotland.org.uk

Appendix A – Glossary of Terms

1. Artificial Intelligence (AI)

- Definition: AI refers to the simulation of human intelligence by machines, particularly computer systems. It involves the ability of machines to perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.

- Example: AI is used in virtual assistants like Siri or Alexa to understand and respond to voice commands.

2. Machine Learning (ML)

- Definition: A subset of AI that involves the use of algorithms and statistical models that enable computers to learn from and make decisions based on data. Unlike traditional programming, where explicit instructions are given, ML allows the system to improve its performance on tasks over time as it is exposed to more data.

- Example: ML is used in recommendation systems, such as those on Netflix, to suggest content based on past viewing behaviour.

3. Deep Learning

- Definition: A type of machine learning that uses neural networks with many layers (hence 'deep') to analyse data and make predictions. Deep learning is particularly effective in tasks like image and speech recognition.

- Example: Deep learning is used in facial recognition technology to identify individuals in photos.

4. Neural Networks

- Definition: A network of algorithms designed to recognize patterns, mimicking the way a human brain operates. Neural networks consist of layers of interconnected 'neurons' that process input data to make decisions or predictions.

- Example: Neural networks are behind the technology that allows self-driving cars to recognize road signs.

5. Large Language Models (LLMs)

- Definition: A type of AI model that has been trained on vast amounts of text data to understand and generate human-like language. These models can perform tasks like text completion, translation, and summarisation.

- Example: ChatGPT is an example of an LLM that can generate human-like text based on a given prompt.

6. Generative AI (GenAI)

- Definition: A branch of AI that focuses on creating new content, such as text, images, music, or even video, based on input data. It can be used to generate creative content or simulate scenarios.

- Example: Generative AI can create realistic images of people who don't exist or generate human-like text conversations.

7. Natural Language Processing (NLP)

- Definition: A field of AI focused on the interaction between computers and humans through natural language. NLP enables machines to understand, interpret, and respond to human language in a way that is both meaningful and useful.

- Example: NLP is used in spam filters to detect and filter out unwanted emails.

8. Algorithm

- Definition: A set of rules or instructions given to an AI system to help it perform a specific task. Algorithms are the building blocks of AI systems, guiding how they process data and make decisions.

- Example: Search engines use algorithms to determine the most relevant results for a user's query.

9. Data Mining

- Definition: The process of analysing large datasets to discover patterns, correlations, and insights that can be used to make informed decisions. It is often a preliminary step in machine learning and AI projects.

- Example: Retail companies use data mining to analyse customer purchase history and recommend products.

10. Training Data

- Definition: The dataset used to teach an AI model how to make decisions or predictions. The quality and quantity of training data significantly impact the effectiveness of the AI model.

- Example: To develop a model that recognises animals in photos, a large dataset of labelled images of animals is used as training data.

11. Bias in AI

- Definition: Refers to the tendency of AI systems to reflect the biases present in their training data, which can lead to unfair or inaccurate outcomes. Bias can be based on gender, race, or other factors, and it is a critical concern in AI development.

- Example: If an AI model is trained predominantly on data from a specific demographic, it may perform poorly on other demographics.

12. Automation

- Definition: The use of AI to perform tasks without human intervention. Automation can streamline processes, reduce errors, and increase efficiency, but it may also rAI'se concerns about job displacement.

- Example: AI-powered chatbots automate customer service by answering common questions without human input.

13. Ethical AI

- Definition: The practice of developing and deploying AI in a way that is fair, transparent, and respectful of human rights. Ethical AI involves considerations like privacy, accountability, and the avoidance of harm.

- Example: Ensuring that an AI system used in recruitment does not discriminate based on race or gender is an aspect of ethical AI.

14. Cognitive Computing

- Definition: A term often used interchangeably with AI, focusing on systems that simulate human thought processes. Cognitive computing systems use AI to solve complex problems by mimicking the way humans think.

- Example: IBM's Watson is an example of cognitive computing used in healthcare to help doctors diagnose diseases.

15. Supervised vs. Unsupervised Learning

- Supervised Learning: A type of machine learning where the model is trained on labelled data, meaning the input data is paired with the correct output. The model learns to make predictions or decisions based on this training.

- Unsupervised Learning: A type of machine learning where the model is given input data without labelled responses. The system tries to learn the patterns and structure from the data itself.

- Example: In supervised learning, an AI model might be trained to recognize spam emails (labelled data), while in unsupervised learning, it might be tasked with grouping similar customer behaviours without pre-labelled outcomes.

16. Prompt

- Definition: In the context of generative AI, a prompt is the input given to an AI model to guide it in producing an output. The prompt can be a question, a statement, or any form of text that instructs the AI on what kind of content to generate. The quality and specificity of the prompt can significantly influence the AI's response.

- Example: If you type 'Write a short story about a cat and a dog' into a generative AI model like ChatGPT, the text 'Write a short story about a cat and a dog' is the prompt.

17. Hallucination (in AI)

- Definition: In the context of AI, particularly generative AI, a hallucination occurs when the AI model produces information or content that is incorrect, nonsensical, or fabricated. This happens when the model generates responses that seem plausible but are not based on real or accurate data.

- Example: If a generative AI model like ChatGPT is asked a question and responds with a confidently stated fact that is completely untrue, that response is considered a hallucination.

Appendix B- References

(Some of these articles are behind a paywall.)

https://medium.com/write-a-catalyst/why-ai-will-never-replace-writers-even-for-years-to-come-904d72284973

https://medium.com/write-a-catalyst/these-9-annoying-words-make-it-obvious-that-your-text-is-written-by-a-i-19d6e163c479

AI timelines: What do experts in artificial intelligence expect for the future? - Our World in Data

https://blog.metamirror.io/gpt-still-understands-nothing-3eed694e00fb

https://medium.com/enrique-dans/lets-use-ai-to-liberate-students-and-create-a-personalized-learning-process-1fb1228d9428

https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

[Global audiences suspicious of AI-powered newsrooms, report finds | Reuters](#)

[An Infinite Shortage of Managers - by Edward Saperia (civicai.uk)](#)

[Panel: Youth Voices & AI - The BIG Survey with Young Scot — Scottish AI Summit](#)

[How to Implement AI — Responsibly (hbr.org)](#)

[(17) Employers flooded with AI resumes | LinkedIn](#)

[AI organisational policies - SCVO](#)

[innovate-us.org/generative-ai-explainer](#)

[https://www.gartner.com/en/articles/when-not-to-use-generative-ai](#)

[https://kozyrkov.medium.com/start-with-why-ai-281391b27e39](#)

[https://pub.towardsai.net/do-not-use-llm-or-generative-ai-for-these-use-cases-a819ae2d9779](#)

[https://albertoromgar.medium.com/generative-ai-is-not-a-magic-wand-45a08d8dbaea](#)

[ChatGPT, AI - Explore - Google Trends](#)

[Statement on the Nature and Purpose of Youth Work (youthlink.scot)](#)

[https://www.gov.uk/government/news/safety-of-advanced-ai-under-the-spotlight-in-first-ever-independent-international-scientific-report](#)

[https://medium.com/geekculture/the-mass-adoption-of-ai-a-new-sigmoid-curve-c86cffe1617](#)

[https://medium.com/the-generator/ai-and-the-decline-of-human-intelligence-eca8d9e651d5](#)