

Artificial Intelligence at the University of Edinburgh



Vice Principal Gavin McLachlan
CIO and Librarian to the University of Edinburgh



AI Background

- ▶ University of Edinburgh - world-leading in AI research and education
- ▶ Emergence of ‘broad’, universally available AI
- ▶ AI has the potential to transform Higher Education
- ▶ Understanding implications will be a **major, long-term undertaking** any opportunities
- ▶ **Risks** will dominate many internal and external discussion discussions
- ▶ **Multi-horizon approach** needed to address immediate and longer-term issues
- ▶ **Participatory and consultative co-creation** must be applied



What is AI?

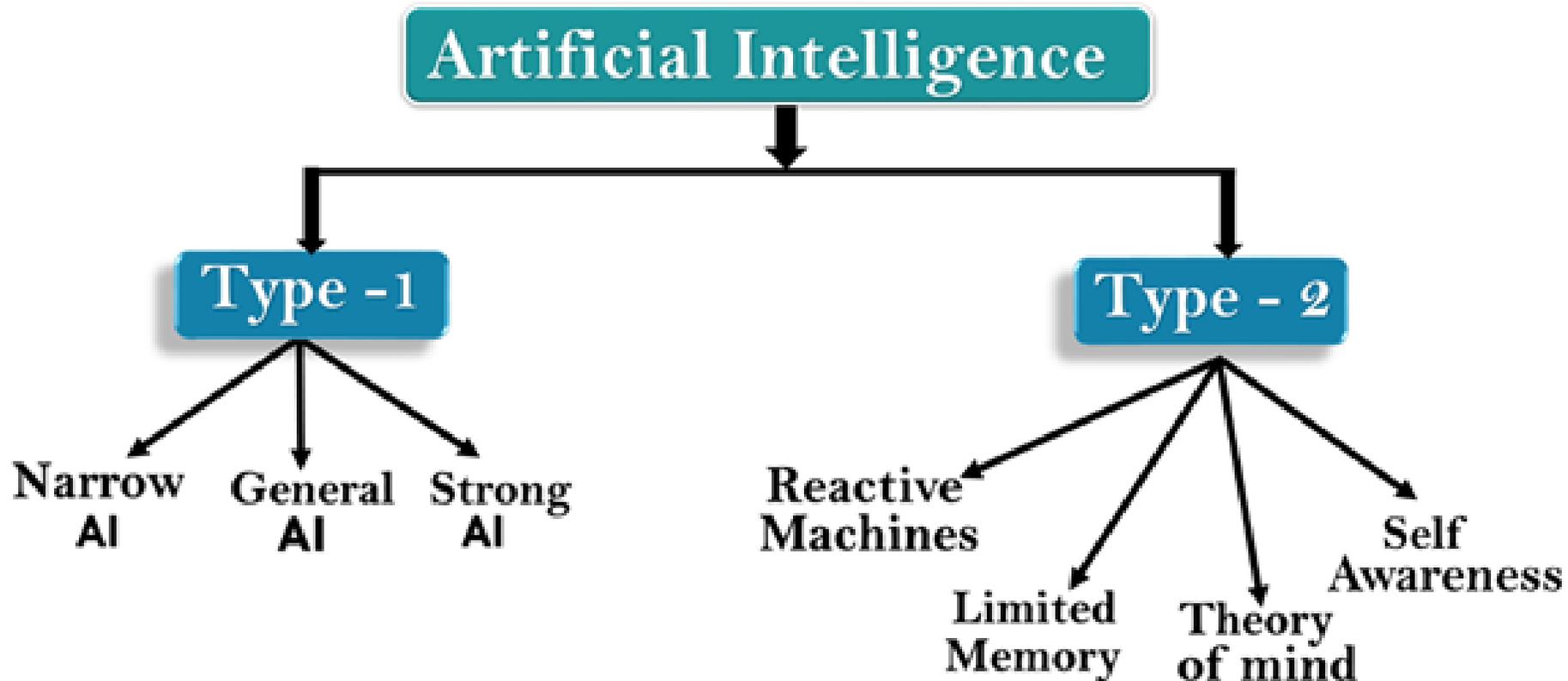
AI systems are computer systems that perform tasks which would otherwise be carried out by humans, and where these tasks involve (1) automated decision making or decision support provided to human users and/or (2) the creation of artefacts normally produced by humans (text, imagery, etc).

AI systems perform these tasks by using complex data-driven algorithms in ways that cannot usually be fully scrutinised by a human operator, because the amount of data exceeds human abilities and/or the underlying computational structures are too large or complex to interpret.

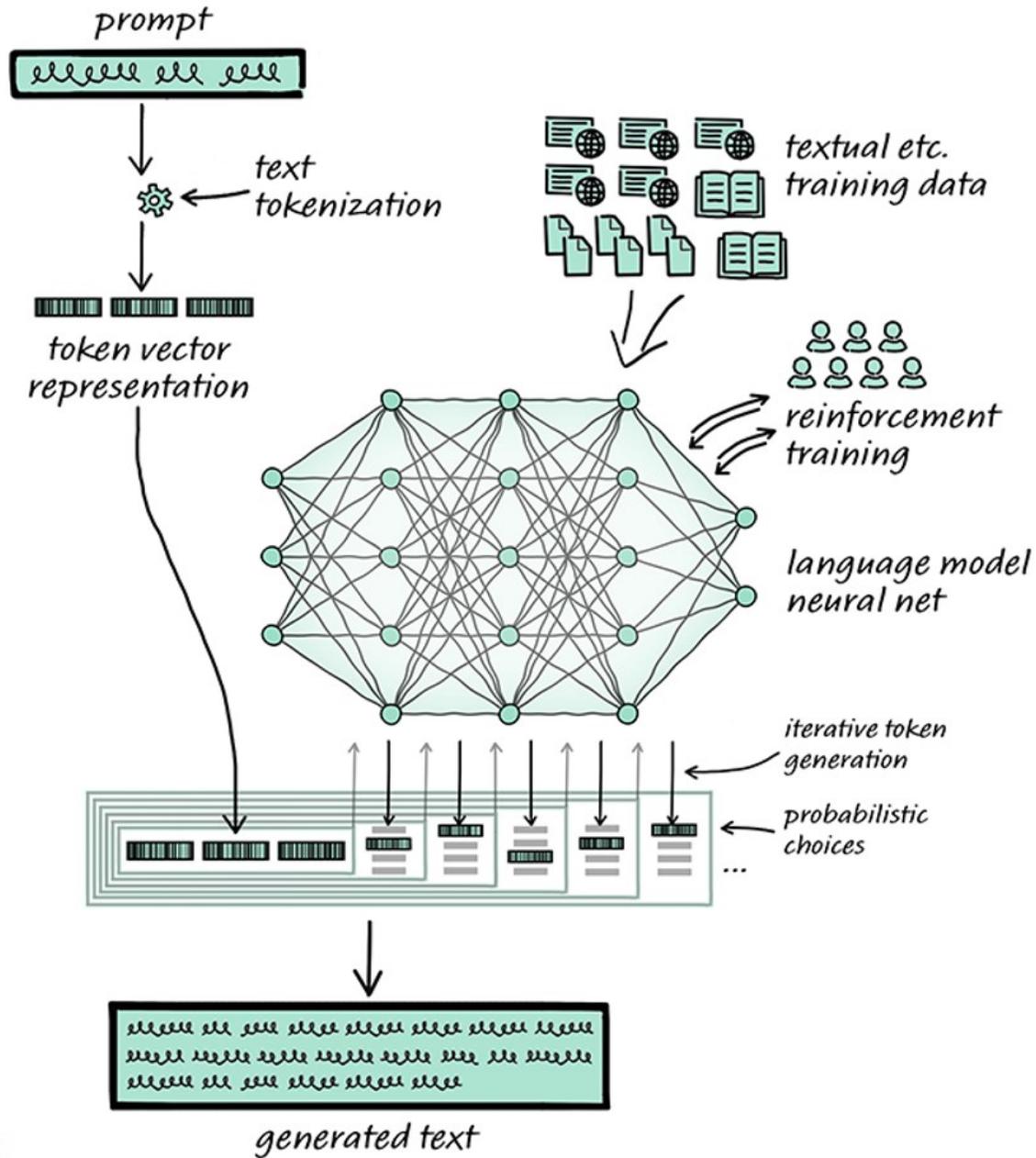


Types of AIs - many types

- ▶ Narrow AIs for a specific purpose: Like Siri, Alexa, facial recognition, DeepMind
- ▶ Generative AIs - example: ChatGPT, Bard, Claude, AI image and video generators
- ▶ Named Entity Recognition - example EDINA automated systematic review



ChatGPT



Machine Learning



Narrow Artificial Intelligence (ANI)

Stage One: Machines imitate human behavior, specializing in one area to solve a problem.

i.e. Siri, ChatGPT, Alexa

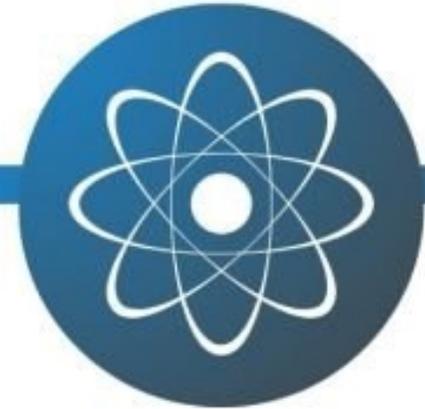
Machine Intelligence



Artificial General Intelligence (AGI)

Stage Two: Machines can continuously learn and are as smart as humans.

Machine Consciousness

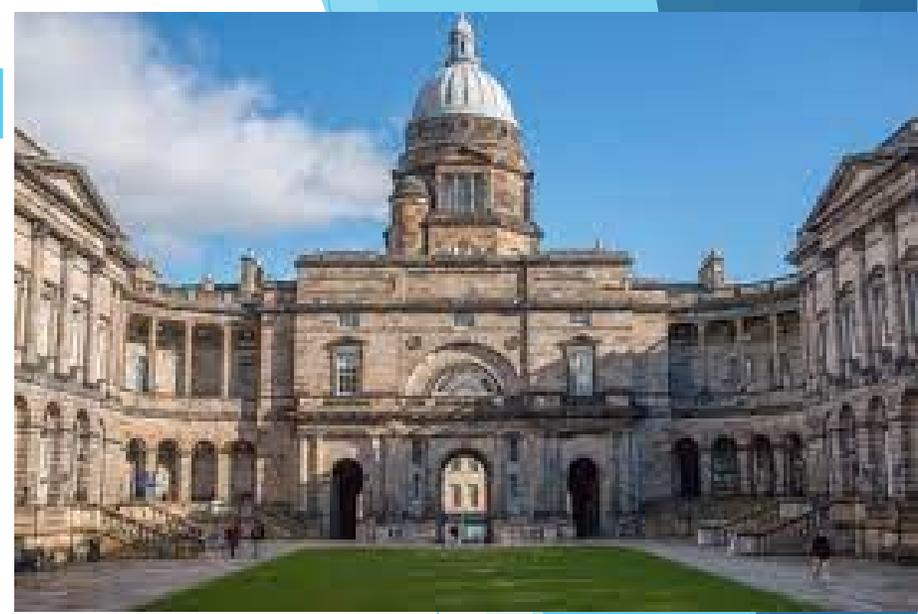


Artificial Super Intelligence (ASI)

Stage Three: Machines that are smarter than humans across the board.

High Educational Sectoral background

- ▶ Existing pressures may result in **resistance to innovation**
- ▶ **Low IT and data maturity** is a common roadblock
- ▶ Structural impediments to **collaboration** between IT providers and users
- ▶ Widespread **lack of understanding** of use of AI and its impacts
- ▶ Huge variations in terms of **digital skills and access**
- ▶ Concerns around championing AI merely as a **‘productivity enhancement’**
- ▶ Gap between **lived student experience** and strategy
- ▶ **Anxieties** around possible AI futures (impact on jobs, learning, science)



Sectoral background

- ▶ Most institutions are at **similar (early) stages of thinking**
- ▶ Government bodies, HE networks, researchers **actively engaged** with the topic (LERU, EU Commission, EU Horizon, ...)
- ▶ **Rapidly evolving technology** and AI economy - hard to keep up
- ▶ Policy lags behind technological developments, this might change in 1-2 years
- ▶ **Fragmented global landscape**, many different policy approaches
- ▶ Much **speculation** against clear and present impacts we already witness
- ▶ **Uncertain long-term futures** for HE, there is little ‘science’ around this



Opportunities & Threats

- ▶ Rapid advances prompted us to consider risks and opportunities
- ▶ Too many or too few AI initiatives?
- ▶ Balance of innovation and understanding/management
- ▶ Risks around ethics, guardrails, climate, legislation, internal standards, duplication of efforts, costs, accessibility, equality of access, sustainability, ...
- ▶ AI equality/parity - We must not create another new dimension of inequality - the AI divide.



What we did at Edinburgh

- ▶ First - decided to embrace AI. Put into our Digital Strategy.
- ▶ Leveraged our existing AI and Data Ethics Board - chaired by Professor Shannon Vallor
- ▶ Published AI guidance for staff and students - June 2023
<https://www.ed.ac.uk/bayes/ai-guidance-for-staff-and-students>
- ▶ Formed an AI Short Life Working Group chaired by Professor Michael Rovatsos. Broad group of experts and stakeholders including academics across wide range of disciplines (AI, Informatics, Data, Law, Politics, Economics, Philosophy/Ethics, Divinity, Climate/Environment, Health), students and IT & Professional Services. No Exec members. Advised on the internal adoption of AI
 - ▶ Brief covered teaching, research, and operations
 - ▶ Built on previous work of AI ethics principles and AI Strategy



AI Working Group Recommendations

	Managing Risks	Developing Capabilities	Creating Opportunities
<i>time</i> ↓	Safeguarding the integrity of assessment methods	Catalogue existing internal AI adoption projects	Pilot a small number of internal AI adoption projects in specific areas
	Governance around AI functionalities in 3rd party SW	Create a central AI communications hub and community of practice	Conduct research on the impact of AI on University communities
	Regulating the use of University data by 3rd party AI tools	Appoint a SRO for AI and additional support roles	Integrate basic AI awareness across our curriculum and PS training
	Provide AI awareness training to managers, staff & students	Develop and implement an AI Impact Assessment process	Develop a long-term strategy for internal AI adoption
	Establish a risk register for AI-related risks to corporate functions		

Edinburgh Language Model (ELM)

Edinburgh's cost effective and supported AI platform for general use and innovation

▶ **ELM Platform – general AI LLM usage:**

- ▶ Web UI providing internal free-at-the-point-of-use access to OpenAI's GPT 3.5 turbo and 4.0 LLMs
- ▶ Uses OpenAI's API, paid for centrally on token basis. Reduces costs by 90%
- ▶ University authentication – control access
- ▶ Comes with set of exemplar prompts, prompt engineering and chats saved. Not chat, prompt or data storage with commercial providers – only secure university storage.
- ▶ Help guidance
- ▶ Zero data retention contract with OpenAI & Microsoft.
- ▶ File upload for chat-like Q&A of file content
- ▶ Moderation violations logged and actioned by the University service desk

ELM

The screenshot shows the ELM interface. At the top left is a hamburger menu icon. Next to it is a dropdown menu labeled "Prompts (Basic Prompt)". A dropdown menu is open, showing options: "Basic Prompt" (selected with a checkmark), "Standard", "Maths Teacher", "Statistician", and "Scientific Data Visualizer". To the right of the prompts menu are buttons for "Upload Data", "Settings", and a help icon (a question mark in a circle). Below these is a "Stop Generating" button and a "New Chat" button. At the bottom is a text input field with the placeholder "Send message" and a "Submit" button. Below the input field is a disclaimer: "ELM can make mistakes. Consider checking any important information." and a "Feedback?" link.

Settings

Model: This setting allows you to choose the large language model that you want to use. The main difference is that using 3.5 is faster and using 4.0 will give you better answers.

Model
GPT 4

Temperature: What sampling temperature to use, between 0 and 1. Higher values like 0.8 will make the output more random, while lower values like 0.2 will make it more focused and deterministic. We generally recommend altering this or top_p but not both.

Temperature*
1

Help [↗](#)

- [Guide to Writing Prompts](#) [↗](#)
- [Glossary of terms](#) [↗](#)
- [Terms of Use](#) [↗](#)
- [Acceptable Use Policy](#) [↗](#)
- [Privacy Notice](#) [↗](#)
- [Cookie Policy](#) [↗](#)
- [Feedback](#) [↗](#)

Edinburgh Language Model (ELM)

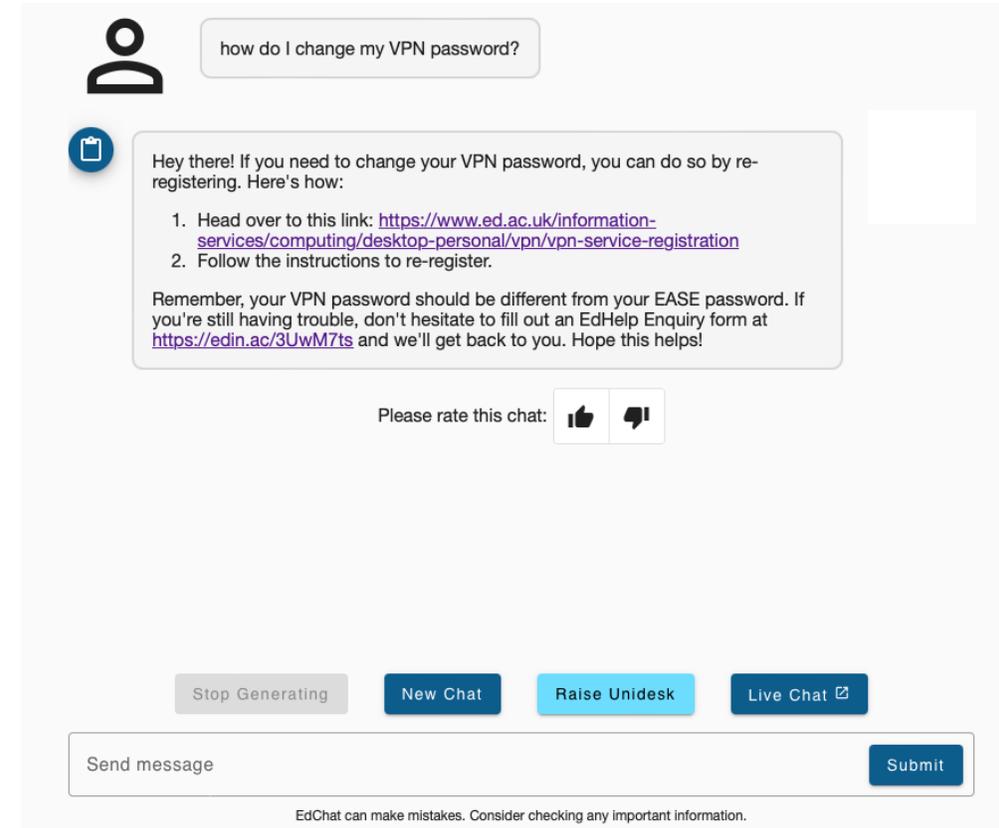
Edinburgh's cost effective and supported AI platform for general use and innovation

▶ **ELM Platform – Innovation platform**

- ▶ Architecture to allow future access to additional (Opensource) LLMs & domain plugins. Llama and others. Ethical, cost and climate LLM alternatives.
- ▶ Chatbots framework - EdHelp student Chatbot
- ▶ Innovation platform for students, researchers and staff – partnership with Informatics and other UoE schools
- ▶ Control our own interface and access for: ethical, monitoring & support, parity of access, accessibility, innovation, climate impact, costs, speed of change/agility.
- ▶ **Ethically approved.** AIDE board review. EQIA, DPIA, Accessibility, Information Security, terms of usage, mandatory user awareness, Zero data retention contract with MS OpenAI. Approved through ISG Ethics board.

Ask EdHelp AI Chatbot

- Restricted to providing info from local knowledge base
- Admin interface for updating knowledge base
- Moderation violation
 - Chat ends
 - Msg depends upon OpenAI moderation violation category
 - Chat pushed to dedicated mlist & service desk category
 - All chats stored in service desk
- Operator request pushed to MS Teams channel
- Overnight updated knowledge items



how do I change my VPN password?

Hey there! If you need to change your VPN password, you can do so by re-registering. Here's how:

1. Head over to this link: <https://www.ed.ac.uk/information-services/computing/desktop-personal/vpn/vpn-service-registration>
2. Follow the instructions to re-register.

Remember, your VPN password should be different from your EASE password. If you're still having trouble, don't hesitate to fill out an EdHelp Enquiry form at <https://edin.ac/3UwM7ts> and we'll get back to you. Hope this helps!

Please rate this chat:

Stop Generating New Chat Raise Unidesk Live Chat

Send message Submit

EdChat can make mistakes. Consider checking any important information.