

Photo by NOAA on Unsplash. See Licence.

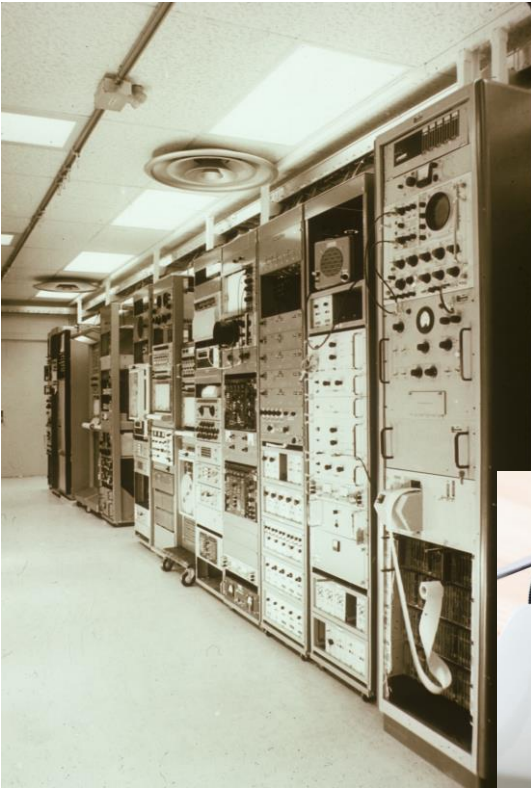


Photo by Markus Winkler on Unsplash. See Licence.

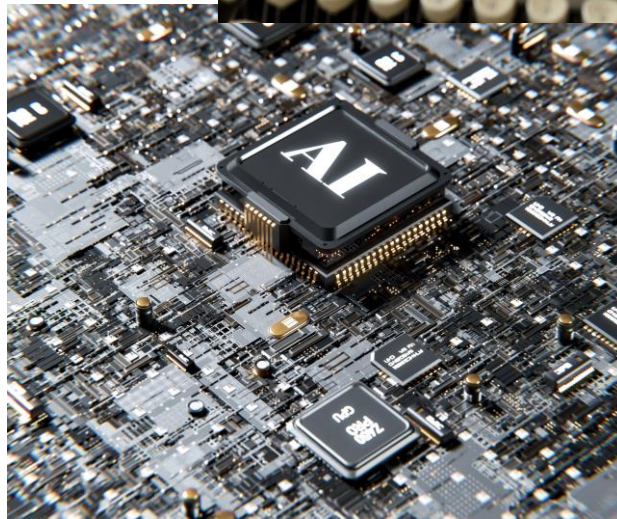


Photo by Igor Omilaev on Unsplash. See Licence.

Guide to Generative AI

Contents



Glossary of Terms

Audience and Goals of this Guide

Setting the Scene – What is Generative AI?

A Digital Skill – Utility in the Workplace

Case Studies in Research

Frequently Asked Questions

More about Artificial Intelligence and Machine Learning

Glossary of Terms

Terms used in this guide may hold differing contexts or meanings between groups of readers within the intended audience. To address potential for ambiguity, terms which may differ in meaning or context are clarified as follows:

- **Digital skill** – The ability to demonstrate a competent level of knowledge, skill, and attitude toward the use of digital products such as software, including programming languages, and digital interfaces to use technology hardware i.e., high performance computing.
- **GPT** – GPT, or “Generative Pre-trained Transformer” is a description of what technology underpins tools such as ChatGPT. “Generative” means that the tool can create new data i.e., a text response to a prompt. “Pre-trained” means that the model has been optimised or learnt from a dataset i.e., Wikipedia pages. “Transformer” is a type of neural network algorithm that is ideal for learning how strings of text are related.
- **Large Language Model** – These models are AI systems which have been trained on vast amounts of human language data. During training, the model learns patterns between words and can eventually reproduce human language when given a prompt.
- **Machine Learning** – This term relates to the creation of systems that learn from “big data” using statistical or mathematical approaches.
- **Model** – A model is a contextual reference to any singular AI system such as OpenAI’s ChatGPT or Microsoft’s Bing Chat Enterprise (also known as CoPilot).
- **Neural Network** – An artificial network of nodes – inspired by the human brain – where inputs (training data) are fed into the network to perform calculations and produce an output (predictions, responses, classifications). Calculations are done in a manner whereby outputs ideally represent patterns in the training data. More computing power allows for more nodes, and more abstract pattern recognition between training data and output data.
- **Prompt** – A description of a task delegated to a generative artificial intelligence system to complete; typically given in text format.
- **Workplace** – The location where you primarily conduct activities to achieve productive outcomes such as an office space, laboratory, or lecture space.

Audience and Goals of this Guide

In line with Griffith University’s eResearch Services value proposition, this guide intends to provide specialist IT support to researchers with respect to generative AI as an emerging technology and increasingly relevant digital skillset.

This guide aims to achieve the following:

- Contextualise positive outcomes and productive benefits associated with generative AI.
- Support the development of digital skills amongst Griffith University researchers.
- Respond to frequently asked questions associated with this emerging technology.

Artificial Intelligence is a field which is undergoing a period of rapid development, where paradigms and long-held beliefs are constantly challenged. In an effort to deliver information which remains current during such rapid development, this guide will only explore fundamental concepts related to generative AI.

For advisory services in relation to a unique generative AI use case, please reach out to eResearch directly at eresearch-support@griffith.edu.au.

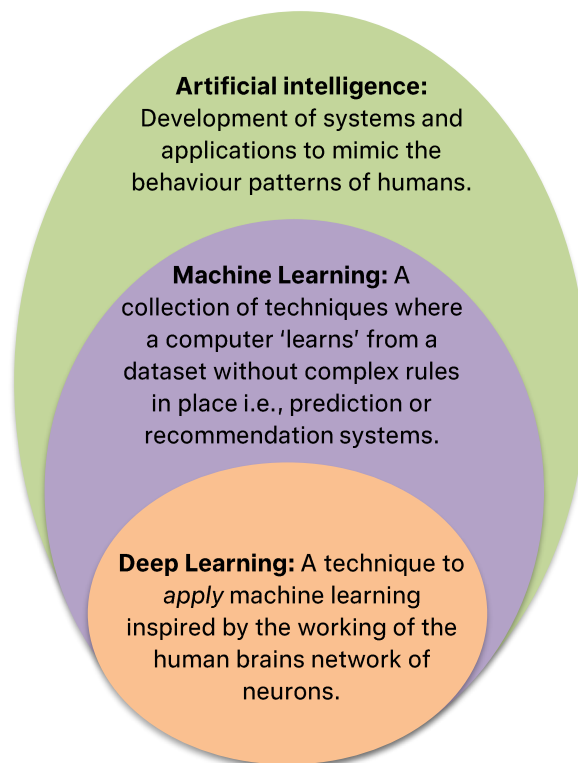
Setting the Scene – What is Generative AI?

Artificial intelligence (AI) is a term used frequently to describe a computing machines capacity to behave with a level of reasoning like that of a human being. Multiple definitions exist for artificial intelligence include these by [IBM](#) and [Google](#). In short, AI can be described as the combination of computer science and robust datasets to create systems which think and act like humans, in solving problems.

Generative Artificial Intelligence (Generative AI, GenAI) is a type of artificial intelligence design to generate new content or data. While popular examples like ChatGPT or Midjourney are focused on generating text or images, Generative AI has much broader applications and has been used to generate everything from new protein structures to simulated medical scans.

While Generative AI has been around for many years, it has recently seen a surge in popularity with the introduction of Large Language Models (LLMs), which among other benefits have opened the ability to interact with Generative AI using natural language, usually in the form of text-based prompts.

Many terms are used alongside AI such as machine learning (ML) and deep learning. The graphic below illustrates the relationships and differences between these terms:



Training data is a critical component of producing any effective generative AI tool; whether it be in the form of text, images, or other. As such, many generative AI models will record or “keep in memory” the prompts it has been given and the output at the time of the prompt to support fine-tuned training of the model.

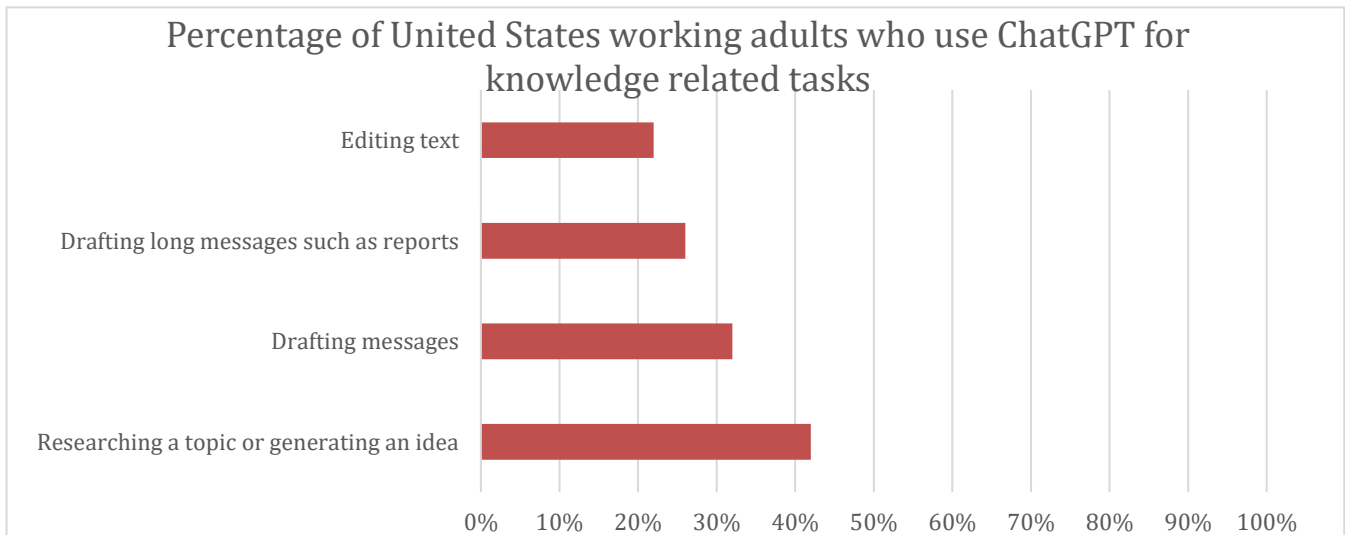
This presents a significant risk when supplying data within a prompt, to a large language model. You can assume that the confidentiality and integrity of any data you feed to the system will be compromised. As such, it is **imperative that no confidential, sensitive information (being information about health, race, sexuality, political views and the like), or other personally identifiable information** is submitted to a large language model. An analysis of the data safety considerations and risks for using LLM's such as ChatGPT can be found [here](#). In summary though you need to remember that the confidentiality of information that you feed to the system will be compromised. This is particularly important to bear in mind if you are dealing with:

- Information that is subject to obligations of confidentiality or valuable trade secrets or industrial know-how;

- Information relating to health or medical research, particularly if it involves patient data where the patient data has not been de-identified so thoroughly that it could never be re-identified or where the patient has consented to use of the data for research purposes but that consent is not broad enough to enable sharing of the data with the AI provider or use of the data for the purposes of training AI (even though that purpose is an ancillary use);
- Information that relates to defence or military-related research or dual use technology (that is not specifically developed for military utilisation but which may have military uses, such as, for example, facial recognition software) or technology which is on the Defence and Strategic Goods List or other technology export controls, in which case it may be a criminal offence to upload the relevant technology to an AI system. For further details see Griffith's Secure Engagement with International Parties Hub at <https://griffitheduau.sharepoint.com/sites/CFI-HUB> and, specifically, the self-assessment checking tool available at <https://griffitheduau.sharepoint.com/sites/CFI-HUB/SitePages/International-Engagement-Checking-Tool.aspx>; and
- Information relating to an invention, in which use of that information may constitute a 'prior use' and prejudice the future patentability (and therefore underlying commercial value) of the invention.

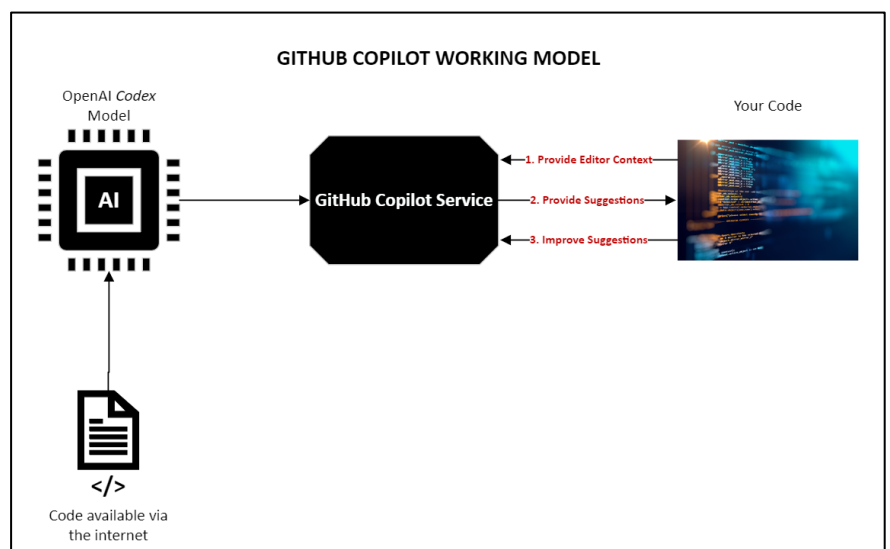
A Digital Skill – Utility in the Workplace

An investment in the order of \$14 billion dollars by Microsoft into OpenAI, the company behind ChatGPT, signals the permanence of this technology in the digital landscape. Furthermore, Microsoft recently stated its intention to integrate large language models (LLMs) into its existing M365 Office suite of applications. Combined, these events highlight a revolution in the ways of working for many professions.



LLMs are not restricted purely to impacting worker productivity in the knowledge space. The development and integration of **GitHub Copilot** is beginning to leverage the capabilities of generative LLMs to support software development and coding productivity.

A study titled *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot⁽¹⁾* assessed developer performance in writing code to achieve a task requiring an intermediate understanding of programming. The study concluded that developer productivity between a group with access to GitHub Copilot and a control group without access, was able to produce code up to 55.8% faster. For researchers writing code and software products of any scale and complexity, the integration of LLMs in developer environments presents itself as a route to achieving net productivity gains.



For any generative AI products, whether it is text, code, or otherwise, care should be taken to verify, test, and validate outputs before publication, production, or final versioning.

Use of AI generated text or code may also constitute a breach of copyright if the AI product has made use of a substantial part of someone else's material in framing its response. Remember that it will be very difficult for you to understand where the AI product has actually sourced its response from, so this becomes a significant element of risk. In circumstances where the text or code is incorporated into publicly available materials (such as academic publications) or included in deliverables for contract research, this risk is increased.

1. Peng, S. et al. (2023) The Impact of AI on Developer Productivity: Evidence from GitHub Copilot [Preprint]. doi:<https://doi.org/10.48550/arXiv.2302.06590>.

Furthermore, if you are using generative AI to generate code, the introduction of the AI generated portions of code into the proprietary elements you have created may lead to significant and, depending on the scale of the AI generated portions, insuperable barriers to proving ownership of the copyright in the final product. This is likely to be more significant an issue in relation to the development of software that may have some commercial application, or a development exercise that is undertaken as a consultancy or commercial research exercise.

As described in the FAQ's below, generative AI is not infallible and exists as a tool to enhance productivity, not replace it.

A primer on prompting

The knowledge related tasks described in the above graph are achieved by submitting **prompts** to LLMs. Prompting can be in the form of posing a general question to an LLM, or posing a question followed by clear instructions in the same prompt. Typical instructional prompts include "Write", "Classify", "Summarise", "Translate", "Order" etc.

Prompts can be supplied iteratively, allowing the LLM to understand more about the problem before providing the final response. This can be an effective technique to reduce probability of incorrect outputs in some situations. For a detailed prompt engineering guide, from fundamentals to more complex use-cases, please see [here](#). As will be reiterated through this document, use great care when submitting prompts to any LLM to ensure no sensitive information is contained within the prompt.

An example of a prompt asking ChatGPT to produce a draft covering letter between researchers for a potential collaboration, is included below. Please note, no personal information (such as name, or workplace) have been included here. Instead, sensitive information has been replaced by generic identifiers such as [person 1], [insert university name] etc. You can see that ChatGPT acknowledges this and continues to shape a response in the correct grammatical context for each placeholder. *There are also typos in the output, emphasising the importance of checking LLM output! You will also note and should be aware of the use of Americanised spelling, which will be off putting for a domestic audience.*

Prompt: Please write a 200-word covering letter, addressed to [person 1] from [person 2] in relation to collaboration prospects on a competitive project application, [insert project name], at their university [insert university name]. These people are peers in the same scientific field and share mutual professional respect for one another, please produce the covering letter that uses a writing tone in line with this fact. Refer to extensive experience [person 2] holds in [insert research field].

Output: [Your Name] [Your Address] [City, State, Zip Code] [Email Address] [Phone Number] [Today's Date]

[Person 1] [Their Position/Title] [University Name] [University Address] [City, State, Zip Code]

Dear [Person 1],

I hope this letter finds you in good health and high spirits. It is with great enthusiasm and anticipation that I write to discuss the exciting collaboration prospects for the competitive project application, [insert project name], at our esteemed university, [insert university name]. Having worked alongside you in the field of [insert research field], I am confident that our combined expertise can lead to groundbreaking advancements.

Your remarkable dedication to research and the strides you have made in [insert research field] have garnered immense respect and admiration from peers and scholars alike. I, too, share the utmost regard for your contributions, which is why I believe that a collaborative effort on [insert project name] holds the potential to create an impact that surpasses our individual achievements.

My own extensive experience in [insert research field], which includes [mention some key accomplishments or research highlights], has provided me with invaluable insights and a profound understanding of the challenges and opportunities that lie ahead. I firmly believe that our collective knowledge, skills, and passion for this project will give us a competitive edge and position us favorably for a successful application.

Collaborating with someone of your caliber is an opportunity I eagerly look forward to. Together, we can synergize our talents to explore uncharted territories and develop innovative solutions that can significantly advance the scientific community's understanding of [insert project focus].

I would be grateful for the chance to further discuss the details of our potential collaboration. Your expertise and input will be invaluable as we shape a compelling project proposal. Please let me know your availability, and I will be more than happy to arrange a meeting at your convenience.

Thank you for considering this proposal, and I am optimistic that our shared vision for excellence will propel us toward success in this endeavor. I eagerly await your response and the prospect of working side by side on this exceptional project.

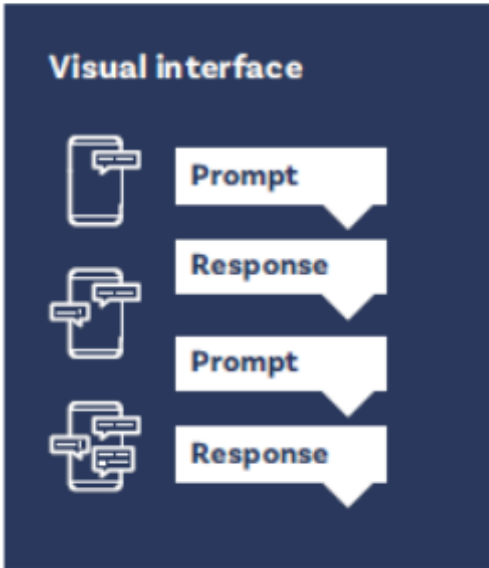
With warm regards and deep respect,

[Your Name]

Case Studies in Research

Before describing case studies and usability scenarios of ChatGPT, a typical LLM user experience is illustrated below¹:

What the user sees



Typical ChatGPT user experience

A user initiates a conversation with ChatGPT by typing a message or question into a chat interface, such as a messaging app or a chatbot on a website.

ChatGPT receives the user's message.

ChatGPT uses its natural language processing capabilities to predict the user's intent from the context of the conversation.

ChatGPT generates a response based on its training and the user's input.

The response is displayed in the chat interface, and the user can respond with follow-up questions or feedback.

The conversation between the user and ChatGPT continues in this way, refining future responses based on the information provided.

What ChatGPT does

ChatGPT selects its responses from a pre-trained Large Language Model (LLM).

An LLM is an AI designed to understand and generate human-like language.

The current model (GPT3.5) has 175 billion parameters and three billion words.

The ChatGPT application shapes its output based on pre-determined rules and previous interactions.

1. Bell, G., Burgess, J., Thomas, J., and Sadiq, S. (2023, March 24). Rapid Response Information Report: Generative AI – language models (LLMs) and multimodal foundation models (MFMs). Australian Council of Learned Academies.

Manuscript Editing and Paper Summaries

In a paper published in *Nature*, two computational biologists constructed a program whereby ChatGPT could edit research papers and produce a revised version of a manuscript according to predefined variables^[1]. Note that while using generative AI to edit research papers saves time, it also adds a very significant degree of uncertainty about who owns the copyright in the output:

Prompt:

"Revise the following paragraph from the introduction of an academic paper (with the title 'An efficient not-only-linear correlation coefficient based on machine learning' and keywords 'correlation coefficient, nonlinear relationships, gene expression') so most of the citations to other academic papers are kept, the text minimizes the use of jargon, the text grammar is correct, spelling errors are fixed, and the text is in active voice and has a clear sentence structure."

Paragraph:

New technologies have vastly improved data collection, generating a deluge of information across different disciplines. This large amount of data provides new opportunities to address unanswered scientific questions, provided we have efficient tools capable of identifying multiple types of underlying patterns. Correlation analysis is an essential statistical technique for discovering relationships between variables [1]. Correlation coefficients are often used in exploratory data mining techniques, such as clustering or community detection algorithms, to compute a similarity value between a pair of objects of interest such as genes [2] or disease-relevant lifestyle factors [3]. Correlation methods are also used in supervised tasks, for example, for feature selection to improve prediction accuracy [4,5]. The Pearson correlation coefficient is ubiquitously deployed across application domains and diverse scientific areas. Thus, even minor and significant improvements in these techniques could have enormous consequences in industry and research.

GPT-3 edits:

~~New technologies have vastly improved data collection, generating a deluge of information across different disciplines. This large amount of data provides new opportunities to address unanswered scientific questions, provided we have efficient tools capable of identifying multiple types of underlying patterns. Correlation analysis is an essential statistical technique for discovering to uncover such relationships between variables [1]. Correlation coefficients are often used in exploratory data mining techniques, such as clustering or community detection algorithms, to compute a calculate the similarity value between a pair of two objects of interest such as , like genes [2] or disease-relevant lifestyle factors related to diseases [3]. Correlation methods They are also used in supervised tasks, for example, for like feature selection, to improve boost prediction accuracy [4,5]. The Pearson correlation coefficient is ubiquitously deployed widely used across many application domains and diverse scientific areas disciplines. Thus Therefore , even minor and significant small improvements in these techniques could this technique can have enormous consequences in a huge impact on industry and research.~~

Final paragraph:

The increasing availability of data has opened up new possibilities for scientific exploration. To take advantage of this, we need efficient tools to identify multiple types of relationships between variables. Correlation analysis is a useful statistical technique to uncover such relationships [1]. Correlation coefficients are often used in data mining techniques, such as clustering or community detection, to calculate the similarity between two objects, like genes [2] or lifestyle factors related to diseases [3]. They are also used in supervised tasks, like feature selection, to boost prediction accuracy [4,5]. The Pearson correlation coefficient is widely used across many application domains and scientific disciplines. Therefore, even small improvements in this technique can have a huge impact on industry and research.

1. Stokel-Walker, C. and Van Noorden, R. (2023) 'What CHATGPT and Generative AI mean for science', *Nature*, 614(7947), pp. 214–216. doi:10.1038/d41586-023-00340-6.

Urban Planning

Urban planning is the design and management of the physical and social aspects of cities and towns. Generative AI has been implemented by planners and designers to increase the efficiency of developing colour coded, domain-specific artefacts to support planning activities.

Through combining existing satellite technology to capture generic imagery of an urban setting, generative AI has been trained by planners to produce an equivalent digital trace to be used in planning workflows^[1]. This technique required an initial outlay of effort by subject matter experts to train an image recognition and generation model. Eventually, the model was capable of reaching an 85% accuracy rate when given unseen satellite imagery to transcribe.



1. O'Malley, J. (2023a, June 14). *Urban planning to disaster relief: How ai-generated maps are helping to improve lives*. The Guardian. <https://www.theguardian.com/location-intelligence/2023/jun/14/urban-planning-to-disaster-relief-how-ai-generated-maps-are-helping-to-improve-lives>

Scripting Low Complexity Code

As mentioned in the previous section, GenAI can be used to quickly script low complexity code. It is important to note that any GenAI produced code should be used with caution, keeping in mind key principles such as bug safety, trust but verify, and open mindedness as described in this article by [Nature](#).

An example of using ChatGPT to script low complexity code, given a prompt of moderate detail is shown below:

JE

Please write a Python script to replace column names with custom values in a data frame with 10 columns. Please mention any libraries required to achieve this, such as Pandas.



Certainly! You can use the `pandas` library in Python to replace column names in a data frame. Here's a Python script that demonstrates how to replace column names with custom values in a data frame with 10 columns:

```
python Copy code

import pandas as pd

# Create a sample data frame with 10 columns
data = {'A': [1, 2, 3],
        'B': [4, 5, 6],
        'C': [7, 8, 9],
        'D': [10, 11, 12],
        'E': [13, 14, 15],
        'F': [16, 17, 18],
        'G': [19, 20, 21],
        'H': [22, 23, 24],
        'I': [25, 26, 27],
        'J': [28, 29, 30]}

df = pd.DataFrame(data)

# Define custom column names
custom_names = ['Column 1', 'Column 2', 'Column 3', 'Column 4', 'Column 5',
                'Column 6', 'Column 7', 'Column 8', 'Column 9', 'Column 10']

# Replace the column names
df.columns = custom_names

# Print the updated data frame
print(df)
```

In this script, we import the `pandas` library using `import pandas as pd`. We then create a sample data frame with 10 columns and define a list of custom column names. The `df.columns = custom_names` line replaces the existing column names with the custom names. Finally, we print the updated data frame.

Make sure you have the `pandas` library installed in your Python environment. You can install it using `pip install pandas` if you haven't done so already.

A study published by Stanford University investigated the question “Do Users Write More Insecure Code with AI Assistants?” to explore the features and limitations of generative AI in developing code. To summarise the study's findings, participants were more likely to introduce security vulnerabilities into their code using GenAI. However, participants who invested time and effort into the creation of prompts eventually led to generating more secure solutions. For the complete paper, please visit [here](#).

Remember though that if you are using generative AI to generate code, the introduction of the AI generated portions of code into the proprietary elements you have created may lead to significant and, depending on the scale of the AI generated portions, insuperable barriers to proving ownership of the copyright in the final product.

Frequently Asked Questions

What are the ethical concerns of generative AI content?

Use of generative AI models and tools come with multiple ethical considerations and concerns, particularly those working with sensitive data:

- **Bias** – An AI model is as only as good as its training data. As such, some model may be trained on datasets which knowingly or unknowingly contain bias toward certain sentiments, views, or conclusions. Subsequently, outputs from a generative AI model trained on datasets with bias may contain the same bias.
- **Misinformation** – Generative AI can be used to produce convincing material (text, image, or video) and if used for malicious intent, can lead to the spread of misinformation. Generative AI is not infallible.
- **Privacy** – Consider that all prompts submitted to a generative AI model are capable of interception and view by a third party. **As such, do not submit personally identifiable information, health or sensitive content in generative AI prompts.**
- **Creative attribution and copyright** – Concerns remain around the attribution of generative AI material. Intellectual property may no longer reside with the user and may be interpreted as attributable to the AI model, or its developers/creators. Alternatively, the outputs of generative AI may not be protected by copyright, as no human author is involved in their creation.
- **Lacking Authentic Expertise** – Generative AI is only as good as its training data, and as such, cannot reason, extrapolate, or think critically outside the bounds of its training data. With a reliance on correct source material, which cannot be guaranteed, **generative AI systems lack the capacity to think independently.** This rigid constraint toward only being capable of synthesising already available information (i.e., the training data) in original formats **should not be misconstrued with the capacity to reason or think independently.**

How can I learn more about the Copyright implications when using generative AI, including referencing of AI output?

Copyright: Under current law, most AI outputs do not have any copyright ownership or protection. For more information regarding AI outputs and copyright, contact the [Information Policy Officer](#).

Referencing AI output: Each style has its own referencing citation for AI outputs. To find the correct AI referencing for your style, see the Library's [Referencing guides](#).

I have heard that generative AI comes with certain legal considerations. Does Griffith University Legal publish advice on this new technology?

Yes, Griffith University Legal offers two published fact sheet materials in relation to the use of AI: Generative AI – General information and Generative AI – Research & Commercialisation Activities. These fact sheets can be found [here](#).

Does generative AI produce misinformation?

Generative AI systems such as ChatGPT have real potential to generate false or misleading information. In addition, since the training data has no **real-time access** to the Internet, i.e., is offline; there is no capacity to query or check alternative sources for validation, confirmation, or updates. Also, there are examples where generative AI models have been shown to create false references to evidence, which do not exist or do not contain the material in question.

Information across all domains updates at a rapid rate, often at a pace too quick to enable effective model training. As such, it is **essential** that generative AI outputs are approached with an emphasis on the mindset: 'trust but verify'. If the information you request from a generative AI system cannot be verified by some means, it is recommended that it is not used.

Why do ChatGPT and other generative AI tools feel life-like?

Generative AI tools are trained on huge datasets specific to the content each tool generates. For example, ChatGPT, being a text based generative tool was trained on 570GB of text databases from across the internet. Similarly, Stable Diffusion, a generative tool that produces images from text input was trained on a dataset containing up to 5 billion image-text pairs.

The capacity for generative AI models to digest enormous amounts of information will mean that any output produced by human prompting will align with large, and very real training inputs. Although these models rely on their training data to generate content, when the model has seen billions of examples of similar content to the human prompt, it becomes easy for the model thus produce a life-like response.

Are generative AI tools like ChatGPT trustworthy?

Although generative AI tools have been trained on huge datasets, they are not always entirely correct in their output. Each system is designed to generate a response that appears most similar to the training data, based on probability, and this probability is never 100%. Generative AI models are prone to present 'hallucination', where the model outputs inaccurate content that is presented as fact. Similarly, AI generated images can look very real, such as this image winning the [2023 Sony World Photography Awards](#).

Generative AI tools, particularly those which generate text, can be used to enhance productivity. However, it is recommended that nothing is asked of a generative AI system which cannot be verified by human means.

Trust, but verify. Inspect what you expect.

If I give a generative AI tool data, where does the data go?

Each generative AI tool will have specific terms of use that govern the rights and responsibilities of both the user and the system developer in handling any data supplied to the generative AI service. As a first point of call, it is always recommended that you familiarise yourself with these terms in the conduct of due diligence.

However, most generative AI tools **will capture inputs** in their entirety for two primary purposes: to further improve the model i.e., using *your input* as training data (**including reservation of future use and sublicensing of inputs**); and monitoring system use for compliance purposes. In addition, some generative AI tool providers may reserve the right to use and sublicense any inputs for any purpose, as determined by the system provider.

Careful consideration should be applied to all data used to prompt a generative AI model. Being sure to remove any sensitive, confidential, or otherwise non-disclosable information prior to the prompt being submitted. Some generative AI service providers have been known to publicly publish datasets to

demonstrate transparency around model training. Finally, many generative AI systems store data outside Australia, which may be enough to contravene Australian privacy and data security laws and regulations.

You need to also remember that the confidentiality of information that you feed to the system will be compromised. This is particularly important to bear in mind if you are dealing with:

- Information that is subject to obligations of confidentiality or valuable trade secrets or industrial know-how;
- Information relating to health or medical research, particularly if it involves patient data where the patient data has not been de-identified so thoroughly that it could never be re-identified or where the patient has consented to use of the data for research purposes but that consent is not broad enough to enable sharing of the data with the AI provider or use of the data for the purposes of training AI (even though that purpose is an ancillary use);
- Information that relates to defence or military-related research or dual use technology (that is not specifically developed for military utilisation but which may have military uses, such as, for example, facial recognition software) or technology which is on the Defence and Strategic Goods List or other technology export controls, in which case it may be a criminal offence to upload the relevant technology to an AI system. For further details see Griffith's Secure Engagement with International Parties Hub at <https://griffitheduau.sharepoint.com/sites/CFI-HUB> and, specifically, the self-assessment checking tool available at <https://griffitheduau.sharepoint.com/sites/CFI-HUB/SitePages/International-Engagement-Checking-Tool.aspx>; and

Information relating to an invention, in which use of that information may constitute a 'prior use' and prejudice the future patentability (and therefore underlying commercial value) of the invention.

How can someone use generative AI to 'test the waters'?

A user can extract an output from ChatGPT using techniques such as prompt engineering, and fine-tuning.

Prompt engineering is the process of crafting input prompts to guide the model toward a specific and focussed output, depending on the context. Examples include rephrasing questions, starting with simple questions, and adding complexity, or building context iteratively over multiple prompt inputs.

Fine tuning is like prompt engineering, except the model is trained on a specific task or domain knowledge base before prompt engineering occurs.

For someone interested in using generative AI in their work, are there reliable resources to help get started?

Multiple introductory short courses are available for those interested in using generative AI in their work:

- [Introduction to Prompt Engineering for Generative AI](#)
- [Introduction to Prompt Engineering for Developers](#) (Note: Experience with programming in Python is highly recommended)
- [Prompt Engineering Guide on GitHub](#)

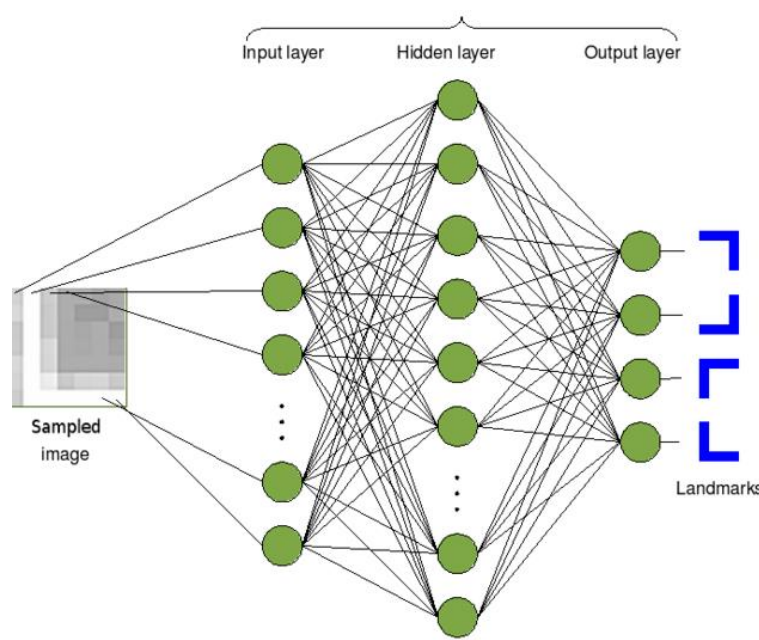
More about Artificial Intelligence and Machine Learning

Generative Artificial Intelligence, generative AI, or GenAI is a sub-type of **artificial intelligence** designed to generate text, video, and image media from a given input called a prompt. Most generative AI systems accept text-based prompts.

AI emerged as a formal field in computer science in 1943, with a formal proposal to build 'artificial neurons' presented by Warren McCulloch and Walters Pitts at the University of Chicago.

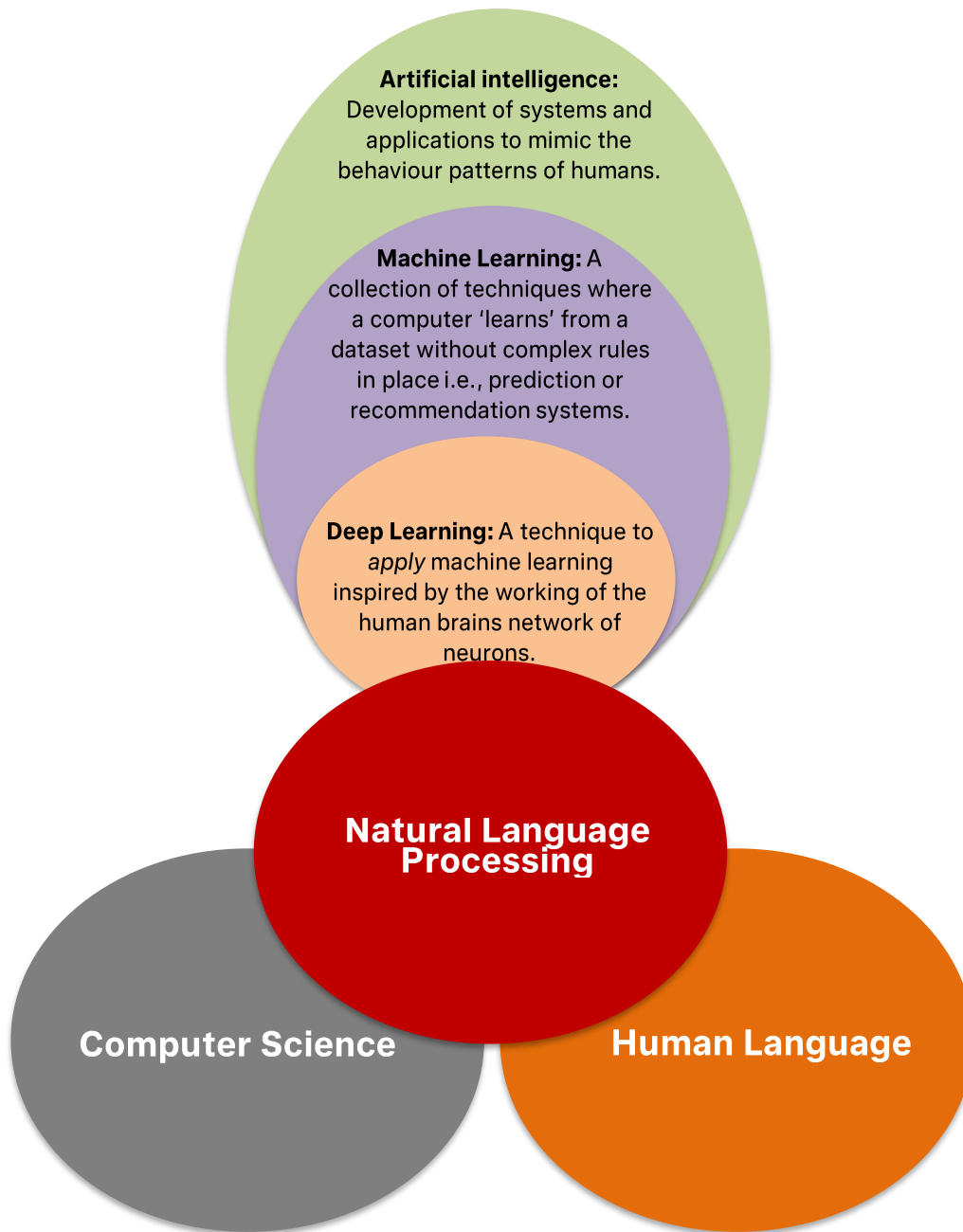
The goal of AI is to develop a synthetic, computational system capable of perceiving and inferring information analogous to the tasks performed by humans. As such, many AI applications relate to tasks which humans already perform, such as: self-driving cars, language comprehension (Siri, Alexa), automated decision-making, and creative material generation (ChatGPT, AI art tools). As use cases become increasingly complex, the data used to **train** a model grows.

Popular generative AI tools such as ChatGPT are referred to as a **large language model (LLM)**. An LLM is a form of deep learning and is constructed using a neural network – one input layer to receive input from a user (i.e., a text-based prompt), one or more intermediate layers of arbitrary size for processing input (i.e. learning) and one output layer. An example of a neural network's structure, designed to recognise landmarks in images, is shown below:



For productivity and digital skill purposes, common LLM's tools such as ChatGPT rely on a subset of deep learning, which intersects human linguistics and computer science called **natural language processing**.

Using our definitions of AI, ML, and deep learning on the previous page, we can include human language and computer science in our mental model to appreciate how central natural language processing is to generative AI technology. (see over page)



Natural language processing, as experienced when using tools such as ChatGPT, is an extension of technology potentially familiar to you already. Such as speech-to-text, and rudimentary customer service chatbots. The generative nature of these existing technologies is where the evolution exists, and the potential for increased productivity and efficiency. More information can be found on natural language processing, [here](#).

