Exhibit A

arts and culture, austerity and anarchism

Language Models as Tools for Thought

During this module you will be given the option to experiment in class with a Large Language Model (LLM) for text generation. This is offered as a *tool for thought*, to be used to explore new ideas about your project through discussion with the primary studio figure, Patrick Geddes.

This exercise is an experiment in teaching and studio discussion, which we hope will offer the following opportunities:

- a better understanding of how language models, AI and 'prompts' work and behave
- an additional method of engaging with class material and historical references
- an exploration of 'critical thinking' and inspiration through writing, extended vocabularies and unexpected associations of ideas
- further critique of your project and new references to research

Note that all discussions with Patrick will be recorded in a spreadsheet and shared with studio members and staff, allowing us to discuss the results together.

BCU publishes clear guidelines on the use of AI and submissions. Any outcomes from this exercise should be openly acknowledged and referenced in your work. Please review the guidelines and discuss your intended uses with your tutor.

The following notes introduce language models and offer a conceptual framework intended to improve your understanding of their behaviour. You may wish to research this topic further and explore discussions regarding prompting techniques as a skill for future interactions with our re-animated historical figure!

---

1 - Introduction to Language Models

Language Models, particularly Large Language Models (LLMs), are not intelligent agents but rather "cultural technologies" that process and generate human-like text. Imagine them as a "haunted compost of books."[1] This evocative metaphor suggests that LLMs are a decomposed, fermented mass of human knowledge, teeming with the ghosts of countless texts.

Ted Chiang's analogy of LLMs as a "blurry JPEG of the web" further illuminates this concept[2]. Just as a compressed image loses fidelity but retains overall structure, LLMs capture the essence of human knowledge while losing precise details.

> *"We use so many different metaphors to talk about LLMs. They have been called "autocomplete on steroids" and "stochastic parrots." The science fiction writer Ted Chiang called ChatGPT a "blurry JPEG of the web." Alison calls them "cultural technologies." Others use metaphors that have more to do with agency and understanding. I'm not sure we've come up with the right metaphor because in some sense, LLMs are all these things."*
>
> *— Melanie Mitchell*

---

2 - How LLMs Work

At their core, LLMs are built on neural networks, which can be understood as complex, multi-dimensional landscapes of information. These networks are incredibly powerful because they can approximate almost any pattern or relationship in data, a property known as universal approximation. In practice, LLMs use multiple layers of these networks, creating a structure that processes language in increasingly abstract ways.

This layered structure allows the model to capture intricate patterns in text, from simple word associations to more complex linguistic structures.

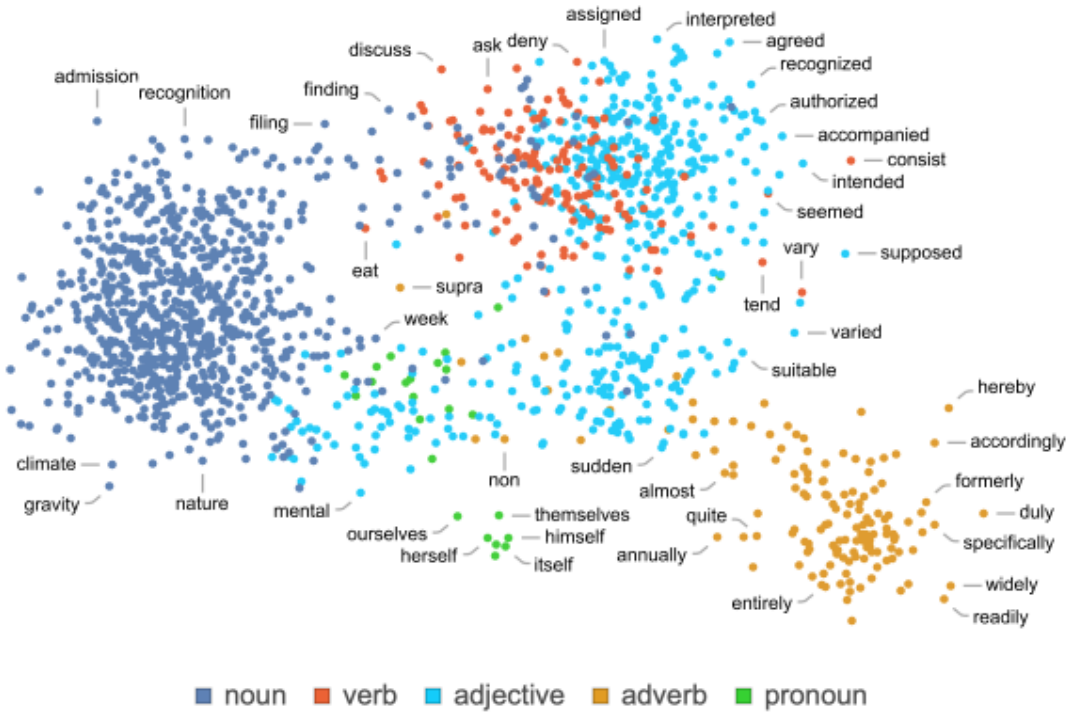

image source: What Is ChatGPT Doing ... and Why Does It Work? — Stephen Wolfram (link)

The training process of an LLM involves exposing the model to vast amounts of text data. As Maggie Appleton explains, this process is akin to "squishing" the structure of language into a mathematical space[3]. The result is a model that can generate human-like text by navigating these complex landscapes of language patterns, without truly understanding the content in the way humans do.

This perspective helps us understand both the power and limitations of LLMs. While they can capture incredibly complex patterns in language, they fundamentally operate on statistical relationships rather than conceptual understanding.

---

3 - LLMs as Tools for Thought

LLMs offer exciting potential as tools for thought, particularly in creative fields like design and architecture. They can serve as collaborative brainstorming partners, offering new perspectives and connections that might not occur to human thinkers alone. LLMs can capture and reframe our collective knowledge, potentially leading to novel insights and ideas.

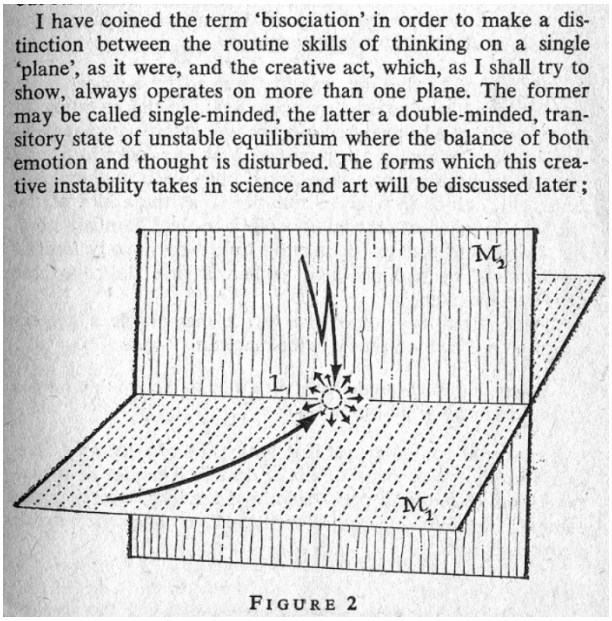
I have coined the term 'bisociation' in order to make a distinction between the routine skills of thinking on a single 'plane', as it were, and the creative act, which, as I shall try to show, always operates on more than one plane. The former may be called single-minded, the latter a double-minded, transitory state of unstable equilibrium where the balance of both emotion and thought is disturbed. The forms which this creative instability takes in science and art will be discussed later;

FIGURE 2

Arthur Koestler's concept of bisociation, introduced in his work "The Act of Creation," involves the combination of unrelated ideas to create new insights. LLMs, with their vast repository of "haunted" knowledge, are uniquely positioned to facilitate bisociative thinking.

By drawing connections between seemingly disparate concepts, LLMs can spark creative leaps that might elude human thinkers. This ability to "bisociate" across vast swathes of information makes LLMs powerful tools for creative problem-solving in fields like architecture and design.

---

4 - LLMs as Cultural Technology

Understanding LLMs as cultural technology places them in a lineage with other transformative innovations like writing and the internet. Alison Gopnik suggests that LLMs could be considered the next step in a series of cultural technologies that have shaped human cognition and society.[4]

- Language
- Writing
- Printing
- Libraries
- Internet
- Wikipedia
- *Large Language Models (LLMs)* ???

Like earlier cultural technologies, LLMs reshape our access to and interaction with information, potentially altering how we think and create. This perspective encourages us to consider LLMs not as isolated AI phenomena, but as part of a broader historical trajectory of human knowledge management and dissemination.

---

5 - Future Potential and Challenges

The future of LLMs in creative fields is both exciting and challenging. As these models evolve, they may become even more powerful tools for ideation and problem-solving. However, their integration into creative processes raises questions about authorship, originality, and the nature of creativity itself. For architects and designers, the challenge will be to harness the potential of LLMs while maintaining ownership to ensure deliberate, mindful, authorial control over the work they create.

---

6 - Limitations and Considerations

While powerful, LLMs have significant limitations. They lack true understanding and can produce convincing but factually incorrect information. As Alison Gopnik and Melanie Mitchell discuss, LLMs don't have the kind of causal understanding that humans develop through interaction with the physical world[5].

In creative fields like architecture, where physical reality is paramount, this limitation is crucial to keep in mind. Ethical considerations, such as copyright issues, model training labour conditions and the potential for bias, also need careful consideration when using AI in creative work.

> *"I like to tell people that everything an LLM says is actually a hallucination. Some of the hallucinations just happen to be true because of the statistics of language and the way we use language. But a big part of what makes us intelligent is our ability to reflect on our own state. We have a sense for how confident we are about our own knowledge. This has been a big problem for LLMs. They have no calibration for how confident they are about each statement they make other than some sense of how probable that statement is in terms of the statistics of language. Without some extra ability to ground what they're saying in the world, they can't really know if something they're saying is true or false."*
>
> *— Melanie Mitchell*

> *"LLMs are not trained to do this. We talk about how they "hallucinate," but hallucination isn't really the right word. To hallucinate would mean that you recognize the difference between the truth and just a bunch of things that people have said. LLMs are not designed to make that distinction."*
>
> *— Alison Gopnik*

footnotes:

1. The phrase "haunted compost of books" was coined by friend of the studio, Dr Justin Pickard.
2. Ted Chiang's analogy is discussed in various articles, including "Chatbots Aren't Human, and We're Making a Mistake if We Think They Are" in The New Yorker.
3. Appleton, Maggie. "Squiggle Theory: Squishing Language into Math." https://maggieappleton.com/squish-structure
4. Gopnik, Alison. This list, showing LLMs as the latest in a series of cultural technologies can be found in several of Gopnik's recent talks or publications on AI and cognitive development often available on YouTube.
5. Gopnik, Alison and Mitchell, Melanie. "How to Raise Your Artificial Intelligence: A Conversation with Alison Gopnik and Melanie Mitchell." https://lareviewofbooks.org/article/how-to-raise-your-artificial-intelligence-a-conversation-with-alison-gopnik-and-melanie-mitchell/