

# Aims

In this assignment, you are required to research the topic of knowledge representation and retrieval in Natural Language Processing (NLP), comparing traditional pipeline-based approaches with modern Large Language Models (LLMs). You will approach this task using two methods: a human-led research process and an AI-assisted process using a Generative AI tool (E.g ChatGPT, Gemini, Claude). You will then critically compare the results and reflect on the implications of using AI Assistance technologies in academic and professional contexts.

## Using Generative AI Tools

For this assignment, you will use a Generative AI tool (e.g., ChatGPT, Gemini, Claude, or similar) to assist with part of your research. You are free to choose any tool you prefer; however, for those unfamiliar with these platforms, ChatGPT is provided here as a baseline example.

If you choose to use ChatGPT, you can access it at <https://chat.openai.com>. You will need to sign up for an account to use the service.

Before using any AI tool, please familiarise yourself with the relevant platform's **Terms of Service** and **Privacy Policy**. For ChatGPT, these can be found at:

- Terms of Use: <https://openai.com/policies/terms-of-use>
- Privacy Policy: <https://openai.com/policies/privacy-policy>

Please ensure that your use of any AI assistant follows these policies and is conducted ethically, responsibly, and in line with academic integrity expectations.



# Task:

In this assignment, you will investigate how NLP systems retrieve and represent knowledge by comparing traditional pipeline-based approaches with Large Language Models (LLMs), such as ChatGPT. You will approach this through both **human-led research** and an **AI-assisted process**.

## Human-led research

Research both **traditional NLP pipelines** and **modern LLM-based approaches** to question answering using academic literature and reputable sources. Summarise how each system stores and retrieves knowledge, and identify the types of data they rely on (e.g. indexed documents, structured sources for traditional systems; pretraining corpora and tokenised representations for LLMs). You may wish to revisit module concepts such as **TF-IDF**, **embeddings**, and **vector similarity**, as well as more recent developments like **attention mechanisms** or **transformer architectures**.

Record and organise evidence of your research process and sources for inclusion as an appendix.

## AI-assisted research

Using only a Generative AI tool (e.g. ChatGPT, Gemini, Claude), investigate both **traditional QA systems** and **LLM-based approaches**. Prompt the tool to explain how each system stores and retrieves knowledge, the types of data involved, and how these approaches differ in practice.

Record and organise your prompts and the full AI responses for inclusion as an appendix.

## Comparison and analysis

Compare and evaluate the results of the human-led and AI-assisted approaches, towards both the traditional QA Systems and LLM-based approaches. Analyse the quality, reliability, and depth of the information provided. Discuss how each approach uses data, and identify any issues such as bias, missing context, or hallucinations. Support your comparison with evidence and examples from both methods.

## Ethical considerations

Assess the potential risks and broader implications of using Generative AI tools in academic and professional contexts. Consider issues such as transparency, academic integrity, bias, and the possible misuse of AI-generated content. Propose recommendations for responsible and ethical use of such tools.

## Reflection


Reflect on your experience completing the assignment. Discuss how your understanding of NLP and Generative AI tools has evolved, including any challenges you encountered when engaging with AI-generated content. Consider how you navigated the boundaries between assistance and authorship, and the strategies you used to maintain academic integrity and ensure the originality of your work.




# Deliverables & Submission Requirements

You must submit a single **PDF report** via Canvas. The report should be a **maximum of 3500 words**, excluding appendices, references, title page, and table of contents.

Your submission must include the following:

- A fully completed **cover sheet**, available in the Assignment Zone on Canvas. This must include all required information, including your full name, student ID, and a statement regarding the use of Artificial Intelligence tools.  
 **Important:** Submissions without a completed cover sheet will **not be marked** and will receive a grade of **0**.
- A **title page** with your full name, student ID, and the final word count of your report.
- A clearly structured **written report** that addresses all elements of the assignment brief.
- A properly formatted **references section**, listing all academic and external sources cited in your report.
- An **appendix** containing:
  - **Evidence of your AI-assisted research**, including prompts and responses (transcripts or screenshots).
  - **Evidence of your human-led research process**, such as notes, article summaries, annotated excerpts, or search history/logs showing how you gathered and evaluated academic sources.

## Formatting and Submission Guidelines:

- Submit your work as a single **PDF document**. Do not submit ZIP files or other archive formats.
-  Failure to submit in PDF format will result in a **10-mark presentation penalty**.
- Word count overages may be subject to penalties in line with University policy: [Programme Regulations and Assessment Policies](#)

## Academic Integrity and Use of AI Tools:

You are expected to interact with a Generative AI tool as part of your research process.

However, the **content of your written report must be entirely your own work**.

Using AI to generate or rephrase sections of your report (beyond providing evidence of its outputs) will be considered **serious academic misconduct**.

Your analysis, comparison, and reflection must demonstrate your own **original thinking and critical engagement** with the topic. The appendix is where AI outputs should appear — not as authored content in your report.



# Help with Referencing

Whenever you need to refer the reader to the source of some information, e.g. a book/journal/academic paper/WWW address, provide a citation at that point within the main body of your report.

**Example 1:** ... as we are all now aware referencing is not trivial (Kendal, 2017)

Provide a reference list towards the end of your research paper (after your conclusions section but before any appendices) that contains:

- References, a list of books/journals/academic papers/URLs etc. that have been directly cited from within the report (see example citation above).
- Any material from which text, diagrams or specific ideas have been used, even if this has been presented in your own words, it must be cited within the main body of the paper and listed in the reference list. It is not enough to list this material in a bibliography.

**Example 2:** For Example 1, (using Harvard system) the reference list would contain the following:

Kendal S., 2017, Referencing standards, International Student Journal, Vol 55, Pages 25 – 30, Scotts Pub., ISBN 1-243567-89

This shows the authors, date published, title of paper (in single quotes), title of journal or conference (in italics), volume, page numbers, and publisher (ISBN desirable but not essential).

For further help see the following book which is available in the library:

- Cite Them Right: The Essential Guide to Referencing and Plagiarism by Richard Pears and Graham Shields

An interactive online version of this guide is available by logging into My Sunderland with your User ID and password and then clicking on Me and Library Resources.

**END OF ASSESSMENT BRIEF**