**Provenance Of Information in Large Language Models (LLMs) using Semantic Similarity Search**

A dissertation submitted to The University of Manchester for the degree of

**Master of Science in Advanced Computer Science in Artificial Intelligence**

in the Faculty of Science and Engineering

**Year of submission**

2025

**Student ID**

10719978

School of Engineering

**Contents**

[Abstract 3](#_Toc207097841)

[Declaration of originality 4](#_Toc207097842)

[Copyright statement 5](#_Toc207097843)

[Acknowledgments 6](#_Toc207097844)

[1. Introduction 7](#_Toc207097845)

[**1.1.** **Background and motivation** 8](#_Toc207097846)

[**1.1.1** **The Foundation of LLMs Architecture** 8](#_Toc207097847)

[**1.1.2** **Generative Pre-trained Transformer (GPT) Models and Training Paradigms** 9](#_Toc207097848)

[**1.1.3** **Information Retrieval (IR) and provenance Tracking** 9](#_Toc207097849)

[**1.1.4** **Legal, Ethical and copyright considerations** 11](#_Toc207097850)

[**1.2 Aims and objectives** 11](#_Toc207097851)

[**1.3 Related work** 11](#_Toc207097852)

[3 Methods 11](#_Toc207097853)

[**3.1 Introduction** 12](#_Toc207097854)

[**3.2 Content** 12](#_Toc207097855)

[**3.2.1 Introduction** 12](#_Toc207097856)

[**3.2.2 Detail** 13](#_Toc207097857)

[**3.2.3 More detail** 13](#_Toc207097858)

[3.2.4 Summary 14](#_Toc207097859)

[**3.3 Summary** 14](#_Toc207097860)

[4 Results and discussion 14](#_Toc207097861)

[**4.1 Introduction** 14](#_Toc207097862)

[**4.2 Content** 15](#_Toc207097863)

[**4.2.1 Introduction** 15](#_Toc207097864)

[**4.2.2 Detail** 15](#_Toc207097865)

[**4.2.3 More detail** 16](#_Toc207097866)

[**4.2.4 Summary** 17](#_Toc207097867)

[**4.3 Summary** 17](#_Toc207097868)

[5 Conclusions and future work 17](#_Toc207097869)

[**5.1 Conclusions** 17](#_Toc207097870)

[**5.2 Future work** 18](#_Toc207097871)

[References 18](#_Toc207097872)

[Appendices 20](#_Toc207097873)

[A Project outline 20](#_Toc207097874)

[B Risk assessment 20](#_Toc207097875)

**Word Count**:

# **Abstract**

This is abstract text.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam

lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum,

erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse

ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient

montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque

cursus luctus mauris.

# **Declaration of originality**

I hereby confirm that no portion of the work referred to in the thesis has been submitted in support

of an application for another degree or qualification of this or any other university or other institute of

learning.

# **Copyright statement**

i The author of this thesis (including any appendices and/or schedules to this thesis) owns cer-

tain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manch-

ester certain rights to use such Copyright, including for administrative purposes.

ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be

made *only* in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and

regulations issued under it or, where appropriate, in accordance with licensing agreements

which the University has from time to time. This page must form part of any such copies made.

iii The ownership of certain Copyright, patents, designs, trademarks and other intellectual prop-

erty (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for ex-

ample graphs and tables (“Reproductions”), which may be described in this thesis, may not be

owned by the author and may be owned by third parties. Such Intellectual Property and Re-

productions cannot and must not be made available for use without the prior written permis-

sion of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv Further information on the conditions under which disclosure, publication and commercial-

isation of this thesis, the Copyright and any Intellectual Property and/or Reproductions de-

scribed in it may take place is available in the University IP Policy (see [http://documents.](http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420)

[manchester.ac.uk/DocuInfo.aspx?DocID=24420](http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420)), in any relevant Thesis restriction decla-

rations deposited in the University Library, The University Library’s regulations (see [http:](http://www.library.manchester.ac.uk/about/regulations/)

[//www.library.manchester.ac.uk/about/regulations/](http://www.library.manchester.ac.uk/about/regulations/)) and in The University’s policy on

Presentation of Theses.

# **Acknowledgments**

Acknowledgments go here.

# **Introduction**

Large Language Models (LLMs) have proliferated in recent years, demonstrating remarkable capabilities across a range of tasks including question answering, text summarisation, and code generation. As these models are trained on vast quantities of data scraped from the internet, the provenance of their output has become a significant concern. Specifically, there are growing concerns regarding potential copyright infringement and the use of unauthorised data in their training corpora. Models such as ChatGPT, Google Gemini, and Llama can produce creative content inspired by various authors and creators, this outputted content does not convey the text of the original works verbatim but rather rephrases or summarises it, contrasts it with other text and generates a result that seems to posses a degree of originality (Riofrio Martinez-Villalba, J.C., 2024).

Tracing information within LLMs is a profound technical challenge. Unlike a traditional database, knowledge in these models is not stored in a searchable format but is instead encoded implicitly across billions of parameters during training, a concept known as parametric knowledge (Yu, H., et al, 2024). This effectively renders the model a "black box," where the internal mechanisms mapping inputs to outputs are opaque. In the context of AI auditing, access to these systems is typically categorized. Black-box access, the most common for proprietary models, permits an auditor only to submit queries and analyse the resulting outputs. In contrast, white-box access provides full transparency into the model's architecture, weights, and training data, allowing for deeper, more direct forms of analysis. Casper, S., et al, (2024) stated that recent evaluations of advanced AI systems have predominantly depended on black-box access wherein auditors can solely query the system and examine the replies. However, white box access to the systems internal operations allows auditors to perform further fine tuning of the models and allows detailed interpretation of the outputs.

While existing research has explored methods like data attribution techniques such as watermarking and influence functions, these methods often focus on the training data's effect on model behaviour rather than pinpointing the direct textual origin of a specific generated output. They can demonstrate that a data point had an influence, but not necessarily what the source text was (Koh, P.W.W., et al, 2019). To address this gap, this paper proposes and evaluates a novel framework for verifying the provenance of LLM generated text using a two step, answer to source matching methodology. First, a standalone LLM is prompted to an answer based on its parametric knowledge of a controlled corpus, in this case, the novel "To Kill a Mockingbird". Second, this generated answer is used as a query in an advanced retrieval system. This system employs semantic search, powered by dense vector embeddings and a cross-encoder re-ranker, to perform a forensic analysis and locate the most probable source passages within the original text.

## **Background and motivation**

This section aims to Provide an overview of the thesis background and related concepts to build understanding of the methodologies and technologies used. The follow up sections will introduce notions related to the foundations of LLMs architectures, GPT models and training paradigms, information retrieval and provenance tracking, legal, ethical, and copyright considerations, and finally the motivation of the project.

### **The Foundation of LLMs Architecture**

LLMs are powered by the transformer architecture moving away from the classic natural language processing (NLP) models such as Recurrent Neural Networks (RNNs). The transformer architecture introduces attention-based models compared to the recurrent mechanisms used by the classic models. Schmidt, R.M., (1912) describes the RNN as a type of neural network architecture used to uncover patterns in sequential data using Multi-Layer Perceptron’s to pass information back in cycle to itself. RNNs process text one word at a time in a sequence, passing information along a chain, while effective for shorter texts, this approach struggles with two major challenges which are the vanishing gradient and the lack of parallelization.

Gradient descent is used to train deep neural networks using backpropagation which initializes parameters randomly and trains repeatedly. These parameters are responsible for computing partial derivatives for the model loss, which either vanishes or explodes as the model’s depth grows due to the chain rule (Abuqaddom, I., et al. 2021). The sequential nature of RNNs means that calculation of a distant word cannot begin until the previous word has been processed, this makes training on massive datasets increasingly slow and computationally inefficient.

The transformer architecture was introduced by Vaswani, A., et al. (2017) in the paper “Attention is all you need”, they proposed a self attention mechanism solving the problem of RNNs. This mechanism allows a model to weigh the influence of all words in the input sequence simultaneously, regardless of their position. Self attention enables a model to capture complex relationships between words and helps it understand context behind a given sequence of words. In a Transformer the model takes in a sequence of tokens and creates three vectors, query , key , and value . Attention is computed by comparing each query with all keys to measure their similarity. These similarity scores are then normalized into weights, which are used to combine the corresponding values into the output representation (Vaswani, A., et al. 2017). Below is the equation:

The other important characteristic of the transformer is the encoder and decoder structures of the architecture. Each is usually constructed using 6 layers of feed-forward networks, normalization, and multi-head attention. While the decoder combines self-attention, which focuses on previously generated tokens, and encoder-decoder attention, which combines encoder outputs with decoder states to help with prediction, the encoder uses multi-head attention to model relationships within the input sequence (Raganato, A. and Tiedemann, J., 2018,).

### **Generative Pre-trained Transformer (GPT) Models and Training Paradigms**

GPT is an outstanding innovation in the field of NLP It is driving researcher’s efforts to create models that are capable of understanding and communicating language in ways that are very similar to those of humans. GPT gained popularity following the release of ChatGPT by OpenAI, a research firm dedicated to creating artificial intelligence (AI) technologies. GPT series has progressed through successive iterations each addressing limitations of earlier NLP approaches that heavily relied on task specific labelled datasets.

Yenduri, G., et al, (2024) explained that GPT-1 was released in 2018, demonstrating the effectiveness of pre-training on large volumes of unlabelled text, using a 12-layer transformer decoder to achieve zero shot performance on many tasks. GPT-2 in 2019 expanded this approach with 1.5 billion parameters, significantly improving performance in tasks such as summarisation and translation through its ability to capture long range dependencies. GPT-3 in 2020 marked a breakthrough with 175 billion parameters trained on the Common Craw corpus enabling it to generate human-like text, perform simple coding, and carry out reasoning tasks. This model was initially updated to GPT-3.5 which included both text and code from diverse web sources. The most recent version GPT-4 is a multi model trained on public and other datasets, featuring extended context windows and reinforcement learning from human feedback. Furthermore, instruction tuning is employed to enhance a model's ability to execute user instructions and perform practical tasks. This is achieved by fine-tuning the model on a diverse dataset of tasks, which have been annotated with human-generated prompts and feedback (Peng, B., et al. 2023).

GPT is trained on a diverse corpus of text data including books, websites and other publicly available sources collected through techniques such as text mining. These sources may include copyrighted works raising concerns about unauthorized use and potential infringement.

### **Information Retrieval (IR) and provenance Tracking**

Information retrieval systems are vital to our everyday online experiences. Their core objective is to identify and return documents from a corpus that are relevant to a user's information need, typically expressed as a query. Two principal paradigms govern the architecture of these systems, their application depending on the nature of the information need, these are traditional lexical models and modern semantic models.

Traditional lexical retrieval also known as sparse retrieval operates on a principle of exact keyword matching. These systems describe documents and queries as high dimensional, sparse vectors, as demonstrated by fundamental algorithms like term frequency-inverse document frequency (TF-IDF) (Sparck Jones, K., 1972) and its successor best match 25 (BM25) (Robertson, S.E. and Walker, S., 1994). Relevance is determined by the similarity of keywords, weighted by their statistical importance. Traditional information retrieval systems determine and retrieve information by computing the relevance between document and query given a document collection and a user query . As can be seen below, this relevance assessment usually uses the similarity metric between document and query (Li, X., et al, 2024).

While highly efficient, these models are inherently limited by the vocabulary mismatch problem whereby they often fail to retrieve relevant documents that use different terminology to describe the same topic.

In contrast semantic retrieval is a dense retrieval which aims to overcome these limitations by matching based on contextual meaning. Li, X., et al, (2024) stated that the deep semantic information of texts is captured by dense retrieval techniques based on the bidirectional encoding representations derived from the BERT model (Devlin, J., et al, 2019), which drastically improves retrieval precision with the recent development of deep learning. The dense retrieval encodes the meaning of queries and documents into dense vector embeddings.

There three key steps that an IR system always follow, Hiemstra, D., (2009) outlined them as the representation of document content, the representation of a user’s information need, and the comparison of these representations to establish relevance. Essentially in a dense retrieval system, a pre-trained LLM encodes documents and query into a document vector and a query vector using an encoder function and then stored in a vector database for indexing and retrieval. In dense retrieval, relevance is determined by calculating the semantic similarity between the query and document vectors, most commonly using cosine similarity.

The final stage of an IR system usually consists of using a cross-encoder model to re-rank the relevant documents retrieved from the first stage of retrieval. Petrov, A.V., et al, (2024) explained that cross-encoder models which encode the query and the documents concurrently as a single textual input, are usually used to provide the best results for document re-ranking. These algorithms are known for being resilient when generalizing across retrieval tasks and domains.

#### Metrics for evaluating IR systems

The effectiveness of an IR system is commonly assessed by evaluation metrics that capture different aspects of the retrieval quality. Precision and recall are most fundamental, with precision measuring a proportion of retrieved documents that are relevant, and recall indicating the proportion of relevant documents retrieved. These measures are frequently integrated into the F1 score to balance the two dimensions. For ranked retrieval tasks, metrics such as Precision@k and Recall@k are used to assess performance at the top of the results list which is especially significant in practical applications where users rarely examine results. More advanced measures including Mean Average Precision (MAP), Normalized Discounted Cumulative Gain (nDCG) and Mean Reciprocal Rnak (MRR), further account for ranking order and graded relevance, providing a more nuanced evaluation of system performance (Manning, C.D., 2008).

### **Legal, Ethical and copyright considerations**

Copyright concerns have become a central issue in the development of LLMs, with several companies facing lawsuits for alleged infringement. Sunstein (2024) reports that The New York Times filed a complaint against OpenAI and Microsoft, claiming extensive copyright violations. The lawsuit argues that OpenAI’s models, including ChatGPT, were trained on millions of the newspaper’s articles without authorization. Evidence presented included more than 100 instances where, when prompted with quotes from Times articles, ChatGPT reproduced substantial portions of the original text with only minor paraphrasing. This case highlights the growing legal and ethical debate around the use of copyrighted material in AI training, raising questions about intellectual property rights, fair use and the boundaries of permissible data usage in machine learning.

## **1.2 Aims and objectives**

The aim of this project is to investigate whether the provenance of LLM’s can be identified by developing and evaluating an information retrieval system capable of linking model responses to their potential source documents.

#### Objectives

Below are the objectives of this project:

* **Design and implement an IR system:** An IR system capable of indexing a representative text corpus such as books or articles and retrieve candidates relative to LLM outputs.
* **Experiment with retrieval strategies:** Sparse retrieval, dense embeddings, and reranking strategies to evaluate how effectively the system can identify likely sources behind LLM generated responses.
* Define and apply evaluation metrics to measure the accuracy and reliability of the provenance tracking system.
* **Result analysis:** To determine the extent to which LLM responses can be traced back to their training data or source corpus.

## **1.3 Related work**

# **3 Methods**

## **3.1 Introduction**

## **3.2 Content**

### **3.2.1 Introduction**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

### **3.2.2 Detail**

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit

sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst.

Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed,

volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui

lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu

ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum

vel, eleifend faucibus, vehicula eu, lacus.

### **3.2.3 More detail**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam

lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum,

erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse

ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient

montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque

cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero.

### 3.2.4 Summary

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

## **3.3 Summary**

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean

faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus

eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula,

urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur

et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est,

nonummy in, fermentum faucibus, egestas vel, odio.

# **4 Results and discussion**

## **4.1 Introduction**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

## **4.2 Content**

### **4.2.1 Introduction**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

### **4.2.2 Detail**

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit

sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst.

Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed,

volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui

lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu

ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum

vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec

odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer

vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer

sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In

hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis

vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non,

adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget

quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris

vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas

sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in

faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend

consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium

ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc

eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque

quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce

sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget,

dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc

placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi

eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum,

eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy

id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa.

Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo

nec dolor.

### **4.2.3 More detail**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam

lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum,

erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse

ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient

### **4.2.4 Summary**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

## **4.3 Summary**

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean

faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus

eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula,

urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur

et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est,

nonummy in, fermentum faucibus, egestas vel, odio.

# **5 Conclusions and future work**

### **5.1 Conclusions**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

### **5.2 Future work**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac,

adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer

id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus

et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem.

Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat.

Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum.

Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor

semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis

quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem

ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus

convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet,

enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae

tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit

ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis

sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in

sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis.

Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui.

Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas.

Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

# **References**

Riofrio Martinez-Villalba, J.C., 2024. Refounding Copyright in the Right to Dialogue: Consequences for ChatGPT and other Artificial Intelligence Services.

Yu, H., Atanasova, P. and Augenstein, I., 2024. Revealing the parametric knowledge of language models: A unified framework for attribution methods. *arXiv preprint arXiv:2404.18655*.

Casper, S., Ezell, C., Siegmann, C., Kolt, N., Curtis, T.L., Bucknall, B., Haupt, A., Wei, K., Scheurer, J., Hobbhahn, M. and Sharkey, L., 2024, June. Black-box access is insufficient for rigorous ai audits. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 2254-2272).

Koh, P.W.W., Ang, K.S., Teo, H. and Liang, P.S., 2019. On the accuracy of influence functions for measuring group effects. *Advances in neural information processing systems*, *32*.

Schmidt, R.M., 1912. Recurrent neural networks (rnns): A gentle introduction and overview (2019). *arXiv preprint arXiv:1912.05911*.

Abuqaddom, I., Mahafzah, B.A. and Faris, H., 2021. Oriented stochastic loss descent algorithm to train very deep multi-layer neural networks without vanishing gradients. *Knowledge-Based Systems*, *230*, p.107391.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, *30*.

Raganato, A. and Tiedemann, J., 2018, November. An analysis of encoder representations in transformer-based machine translation. In *Proceedings of the 2018 EMNLP workshop BlackboxNLP: analyzing and interpreting neural networks for NLP* (pp. 287-297).

Yenduri, G., Ramalingam, M., Selvi, G.C., Supriya, Y., Srivastava, G., Maddikunta, P.K.R., Raj, G.D., Jhaveri, R.H., Prabadevi, B., Wang, W. and Vasilakos, A.V., 2024. Gpt (generative pre-trained transformer)—A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. *IEEE access*, *12*, pp.54608-54649.

Peng, B., Li, C., He, P., Galley, M. and Gao, J., 2023. Instruction tuning with gpt-4. *arXiv preprint arXiv:2304.03277*.

Sparck Jones, K., 1972. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, *28*(1), pp.11-21.

Robertson, S.E. and Walker, S., 1994, July. Some simple effective approximations to the 2-poisson model for probabilistic weighted retrieval. In *SIGIR’94: Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, organised by Dublin City University* (pp. 232-241). London: Springer London.

Li, X., Jin, J., Zhou, Y., Zhang, Y., Zhang, P., Zhu, Y., & Dou, Z., 2024. From Matching to Generation: A Survey on Generative Information Retrieval. *ACM Transactions on Information Systems*, 43, pp. 1 - 62. <https://doi.org/10.1145/3722552>.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2019, June. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).

Hiemstra, D., 2009. Information retrieval models. *Information Retrieval: searching in the 21st Century*, pp.1-19.

Petrov, A.V., MacAvaney, S. and Macdonald, C., 2024, March. Shallow cross-encoders for low-latency retrieval. In *European Conference on Information Retrieval* (pp. 151-166). Cham: Springer Nature Switzerland

Manning, C.D., 2008. *Introduction to information retrieval*. Syngress Publishing,.

Sunstein., 2024, The New York Times V. openai: The biggest IP case ever, Sunstein LLP. Available at: <https://www.sunsteinlaw.com/publications/the-new-york-times-v-openai-the-biggest-ip-case-ever> (Accessed: 26 August 2025).

# **Appendices**

# **A Project outline**

Project outline as submitted at the start of the project is a required appendix. Put here.

# **B Risk assessment**

Risk assessment is a required appendix. Put here.