# A battle between
# A Generative AI Large Language Models (LLM) and a Natural Language Process (NLP).

***Can the unethical use of Gen AI Large Language models***

***be detected and conquered by Natural Language Processes?***

Dr Helga Hambrock
Instructional Designer
Sefako Makgatho Health Sciences University

# Background

- Educators have increased concerns about the use of Generative AI about the temptation of students plagiarising by copying content directly from AI language generating tools and submitting the content as their own (Lee & Kwon, 2024).

- This is a battle that educators are facing especially as the GenAI Large Language Models (LLM) are becoming more sophisticated and can produce astonishing human-like writing (Lee, et, al.,2024).

# AI LLM and NLP
# SECTION 1: COMPARISON

TERMS HISTORY FUNCTIONALITIES

In the first part of the study a literature review is presented which includes a clarification of terms and categories in which two identified AI Language models (the LLM and the NLP) are divided.

Thereafter history of the programs is presented to understand the aim and context in which they were developed.

Finally the functionalities of each are presented.

TEST

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# AI LLM and NLP
# SECTION 2: Test


In the second part a test text that was generated by the generative LLM AI software with additional test added by the researcher is pushed through the NLP and is presented and analysed.

Finally, the positives and negatives of the process are discussed and ends with suggestions from an educator's perspective.

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# TERMS

## Natural Language Processes (NLP)

**NLP** is a subfield of Artificial Intelligence that is focused on enabling computers to understand and process human languages, to get computers closer to a human-level understanding of language.

Machines were initially taught to mimic basic human tasks. As time moved, so did the ambition of those who programmed these machines.

**NLP** models include prediction and text recognition models such as Turnitin.

ps://help.turnitin.com/integrity/investigator/dashboard/prediction-ore.htm#:~:text=It%20is%20calculated%20using%20Turnitin's,human%2Dlevel20understanding%20of%20language.

The prediction score is a score between 0 and 1 that is attributed to each student in the Dashboard. The closer the score is to 1, the more likely it is there is something worth investigating.

TERMS

# Large Learning Models (LLMs)

**Large language models** are also referred to as **neural networks (NNs)**, which are computing systems inspired by the human brain. These neural networks work by using a network of nodes that are layered, much like neurons.

Like the human brain, **large language models** must be **pre-trained and then fine-tuned** so that they can solve text classification, question answering, document summarization, and text generation problems, which is also known as **Generative Artificial Intelligence**. Example- ChatGPT

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# History

## NLP

Early NLP systems were rule-based, relying on sets of hand-coded rules to interpret language. However, the field has evolved.
Modern NLP uses statistical and machine learning techniques, allowing machines to learn language patterns from vast datasets. This learning enables NLP systems to perform tasks like sentiment analysis, language translation, and speech recognition and text recognition.

## LLM- GEN-AI

OpenAI'sGPT 3.5/4 and Google's BERT represented monumental leaps. These models, trained on vast datasets, could not only understand and generate language with unprecedented fluency but also perform a variety of language tasks, from translation to question-answering.

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# History

## Where it all began

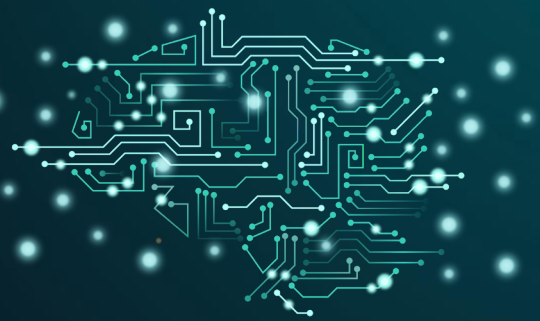
### NLP

1950's

Georgetown-IBM experiment in 1954. **Translation** of 60 Russian sentences into English

1960's

Noam Chomsky introduced the theory of a 'universal grammar' provided a structured approach **to understanding language** – a framework that could be encoded into machines.



### LLM: GEN-AI

Many years passed, and the machines learned to speak more fluently, more like us. They evolved into what we term Large Language Models (LLMs).

ELIZA, a simple chatbot that could mimic human conversation.

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# History

## Where it all began

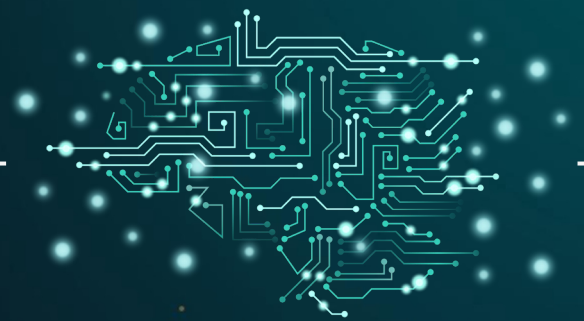
### NLP

1980's 1990's

The field of NLP expanded, incorporating **statistical methods**. This period marked a shift **from rule-based systems to those that learned from large datasets.**

2010's

The '90s were a **renaissance for NLP,** with the advent of **machine learning techniques** that allowed for more nuanced language understanding and generatio



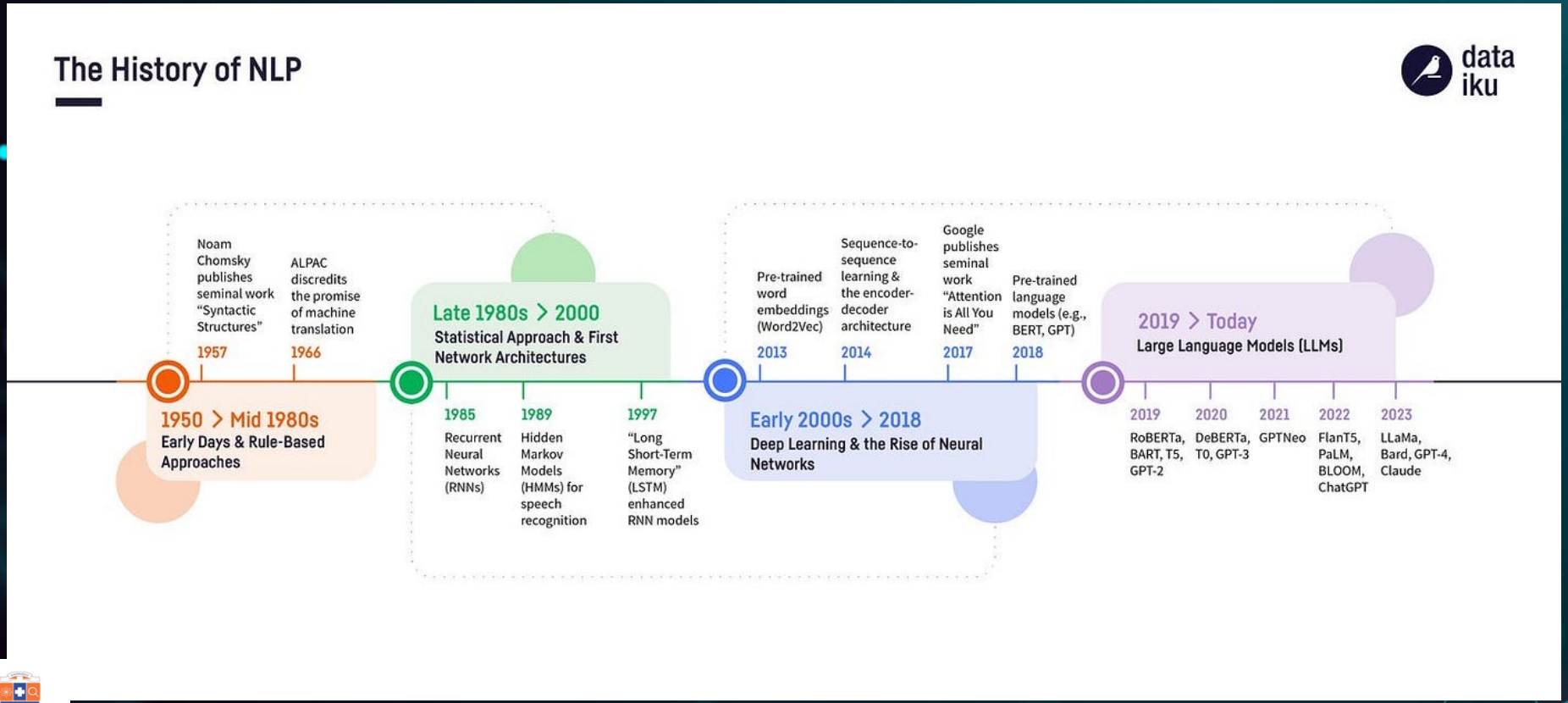
### LLM: GEN-AI

**Generative AI improved by advancements in neural networks and an explosion in computational power.** This era saw the creation of AI models that could **generate realistic images, compose music, and even write coherent pieces of text. 2010**

The introduction of the first neural language model. This model used **neural networks to process and generate language, marking a departure from the rule-based systems of the past.**

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# History


The History of NLP
data iku
1950 > Mid 1980s
Early Days & Rule-Based Approaches
1957
Noam Chomsky publishes seminal work "Syntactic Structures"
1966
ALPAC discredits the promise of machine translation
Late 1980s > 2000
Statistical Approach & First Network Architectures
1985
Recurrent Neural Networks (RNNs)
1989
Hidden Markov Models (HMMs) for speech recognition
1997
"Long Short-Term Memory" (LSTM) enhanced RNN models
Early 2000s > 2018
Deep Learning & the Rise of Neural Networks
2013
Pre-trained word embeddings (Word2Vec)
2014
Sequence-to-sequence learning & the encoder-decoder architecture
2017
Google publishes seminal work "Attention is All You Need"
2018
Pre-trained language models (e.g., BERT, GPT)
2019 > Today
Large Language Models (LLMs)
2019
RoBERTa, BART, T5, GPT-2
2020
DeBERTa, T0, GPT-3
2021
GPTNeo
2022
FlanT5, PaLM, BLOOM, ChatGPT
2023
LLaMa, Bard, GPT-4, Claude

# Functionalities

## BASIC AI LANGUAGE MODELS

Vertiv (2024)

01  ***General language models***

***Predicting*** *the next word or words based on the preceding words in a sentence. WORD, GRAMMARLY*

02  ***A simple probabilistic language model***

*Machine* ***translation, autocorrect, voice recognition****, and autocomplete features - to* ***fill in the following word*** *for the user or suggest possible word sequences. WORD, GRAMMARLY*

03  ***Advanced models, including transformer models***

***Including transformer models****, providing* ***more accurate predictions of the next word by considering surrounding*** *words and context across the entire text.* BERT (Google)

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Functionalities

# GENERATIVE AI LARGE LANGUAGE MODELS

Rashi, et.al; (2024)

***Generative Artificial Intelligence (GenAI) language models***

- GenAI models **recognize patterns** and **"understands" how the data is related and** then **generates content that is unique**.
- This ability of **GenAI** models to **create new content** offers immense potential for revolutionizing the educational landscape.

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# GENERATIVE AI - LARGE LANGUAGE MODELS

Rashi, et.al, (2024)

***Generative Artificial Intelligence (GenAI) language models***

- Teachers can use it for **curriculum design, syllabus writing, teaching, research and administration.**
- For **teaching and learning it also provides constructive feedback at scale, fostering iterative learning and improving writing skills.**
- This ability is especially useful in education for students with learning disabilities, anxiety, or language barriers.
- Furthermore, GenAI helps educational institutions provide customized support and information to students, automate tasks such as scheduling events, and generate promotional content, etc

# GENERATIVE AI - LARGE LANGUAGE MODELS

Rashi, et.al, (2024)

***Generative Artificial Intelligence (GenAI) language models***

- Teachers can use it for **curriculum design, syllabus writing, teaching, research and administration.**
- For **teaching and learning it also provides constructive feedback at scale, fostering iterative learning and improving writing skills.**
- This ability is especially useful in education for students with learning disabilities, anxiety, or language barriers.
- Furthermore, GenAI helps educational institutions provide customized support and information to students, automate tasks such as scheduling events, and generate promotional content, etc

# Functionalities

## NLP

Language Decoded and Refined

Translation

Autocorrect and Autocomplete

Conversational AI

Automated Text Summarization:

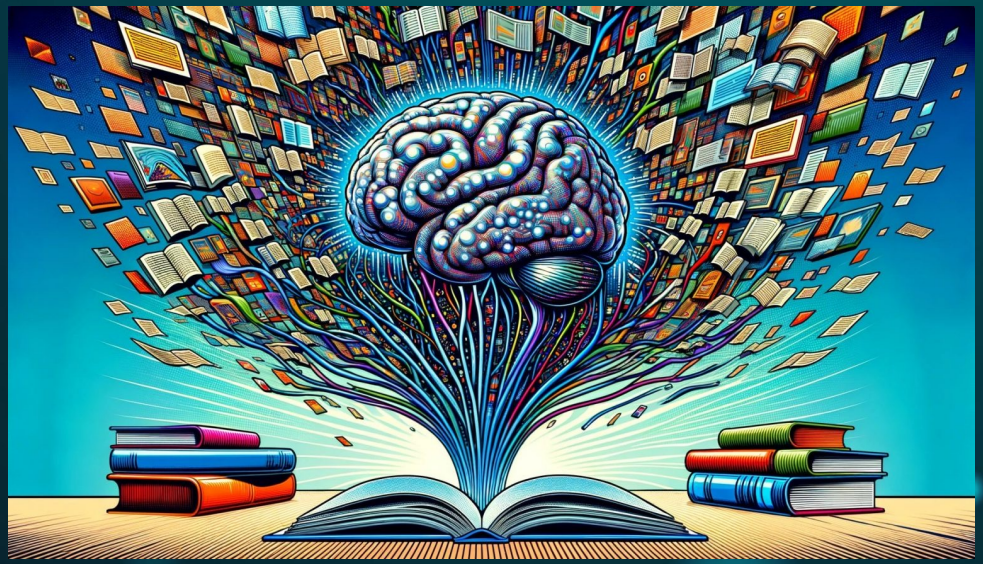
Sentiment Analysis

Email Filters

Speech Recognition

Interactive Learning Tools:

Content Categorization:

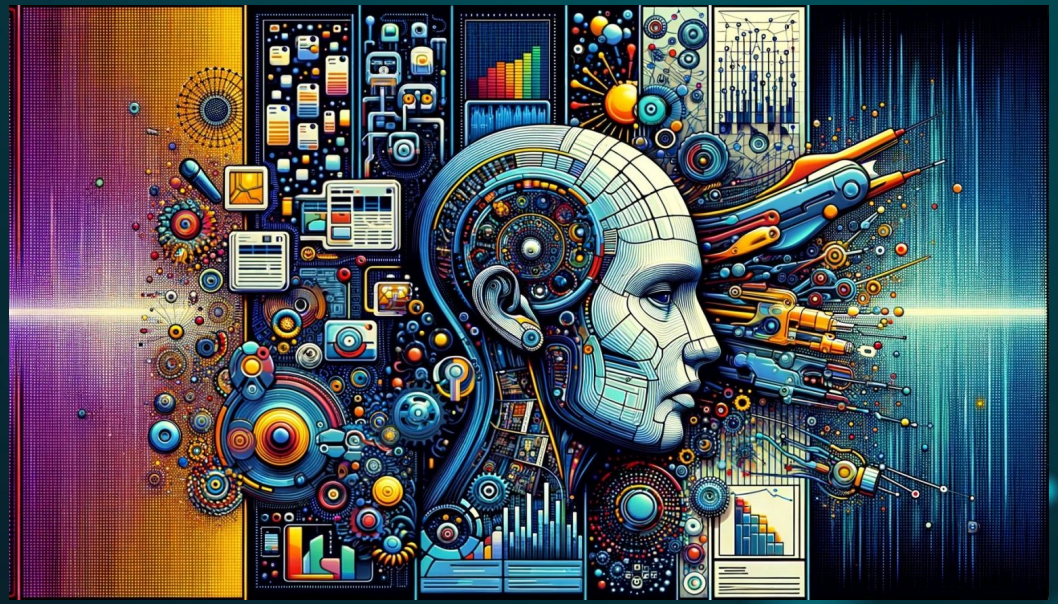


SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Functionalities

## LLMs

- Enhanced Chatbots
- Deeper Sentiment Analysis
- Translation and Localization
- Financial Analysis
- Medical Diagnosis and Research
- Legal Document Analysis



SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Functionalities

## Generative AI

Generative AI Use Cases: Creation and Innovation

- Content Creation
- Automating Workflows
- Business Strategy Development
- Worker Augmentation
- Healthcare Innovations
- Architectural and Engineering Design
- Procedural Content Generation in Gaming
- Data Synthesis for Training AI Models

Each of these technologies, with its unique capabilities, is reshaping how we interact with language, create content, and analyze data.

GenAI vs. LLMs vs. NLP: A Complete Guide (scribbledata.io)

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Processes

## NLP


SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Processes

## LLM


SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Processes

## LLM - NLP



SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

Test

# Gen-AI vs NLP

Methodology

## LLM- INPUT

Request LLM to write an Essay on the advantages and disadvantages of using a LLM (ChatGPT) for essay writing, by students.

## OUTPUT

Essay written by LLM –Gen AI

## HL-INPUT

Thereafter, the researcher added an introduction, a few linking sentences in between and a conclusion to the essay.

## NLP -INPUT

Combined document submitted to NLP.

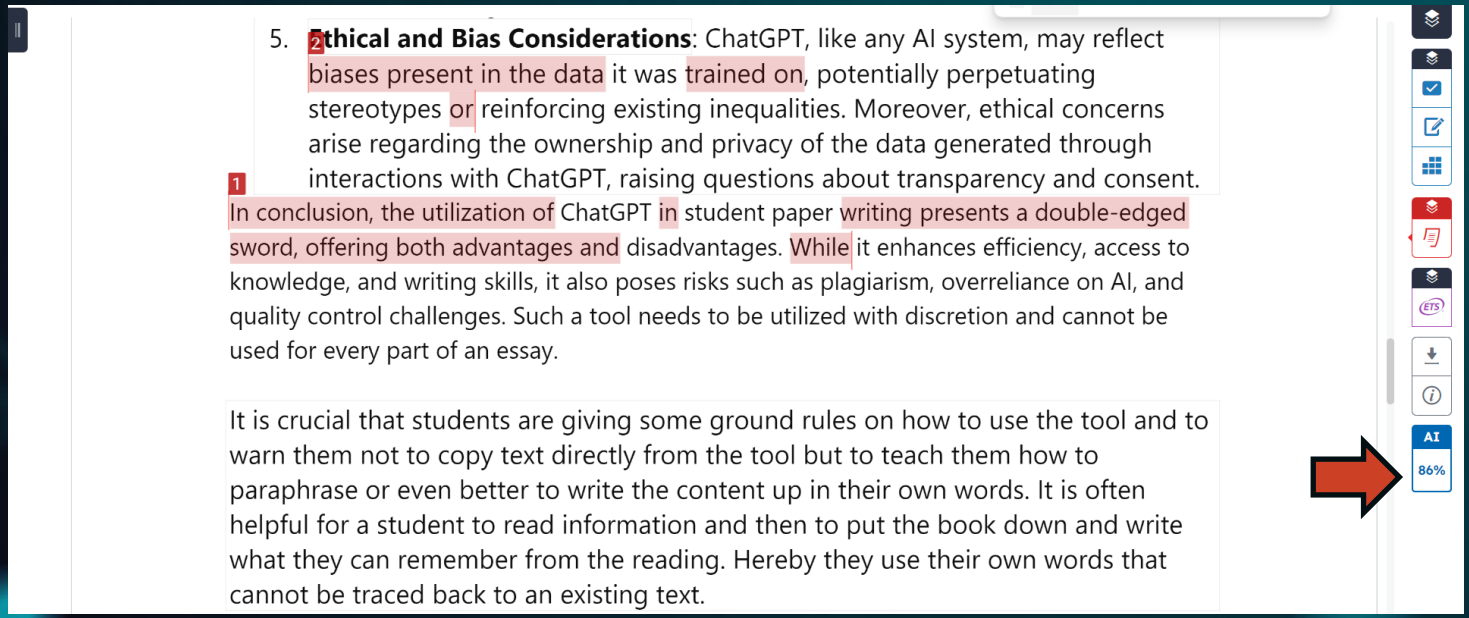
## NLP- OUTPUT

Result

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Result

# Example of AI detection



5. **Ethical and Bias Considerations**: ChatGPT, like any AI system, may reflect biases present in the data it was trained on, potentially perpetuating stereotypes or reinforcing existing inequalities. Moreover, ethical concerns arise regarding the ownership and privacy of the data generated through interactions with ChatGPT, raising questions about transparency and consent.

In conclusion, the utilization of ChatGPT in student paper writing presents a double-edged sword, offering both advantages and disadvantages. While it enhances efficiency, access to knowledge, and writing skills, it also poses risks such as plagiarism, overreliance on AI, and quality control challenges. Such a tool needs to be utilized with discretion and cannot be used for every part of an essay.

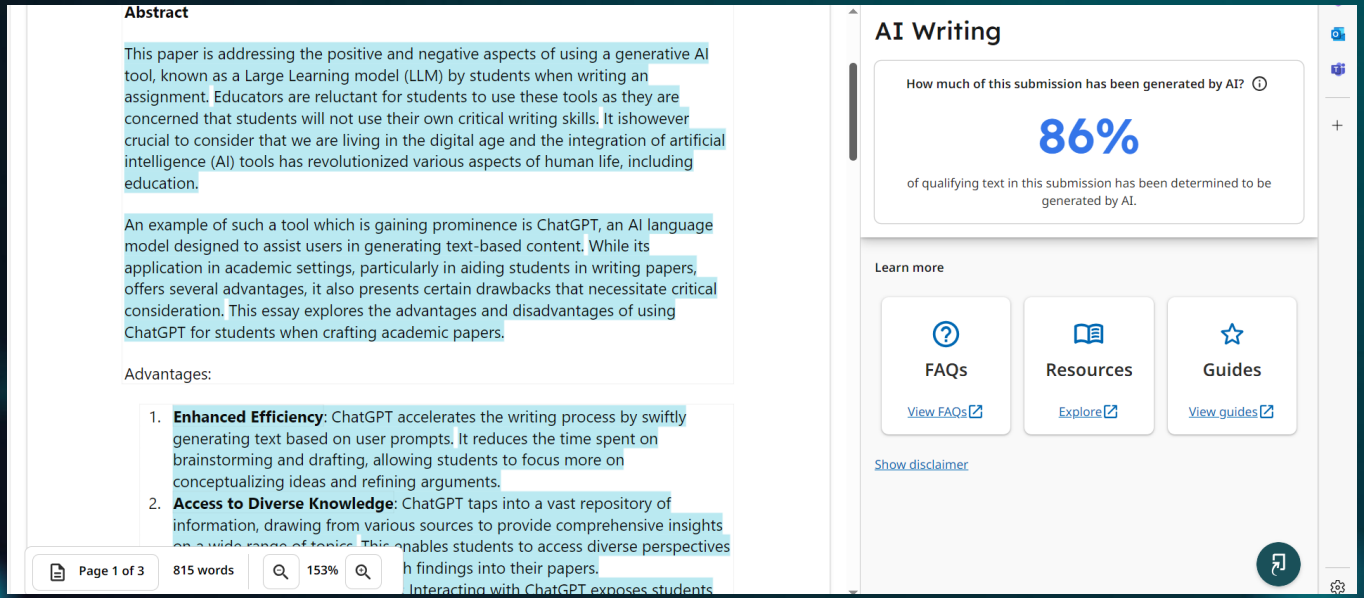
It is crucial that students are giving some ground rules on how to use the tool and to warn them not to copy text directly from the tool but to teach them how to paraphrase or even better to write the content up in their own words. It is often helpful for a student to read information and then to put the book down and write what they can remember from the reading. Hereby they use their own words that cannot be traced back to an existing text.

AI 86%

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Result

# Example of AI detection



**Abstract**

This paper is addressing the positive and negative aspects of using a generative AI tool, known as a Large Learning model (LLM) by students when writing an assignment. Educators are reluctant for students to use these tools as they are concerned that students will not use their own critical writing skills. It ishowever crucial to consider that we are living in the digital age and the integration of artificial intelligence (AI) tools has revolutionized various aspects of human life, including education.

An example of such a tool which is gaining prominence is ChatGPT, an AI language model designed to assist users in generating text-based content. While its application in academic settings, particularly in aiding students in writing papers, offers several advantages, it also presents certain drawbacks that necessitate critical consideration. This essay explores the advantages and disadvantages of using ChatGPT for students when crafting academic papers.

Advantages:

1. **Enhanced Efficiency**: ChatGPT accelerates the writing process by swiftly generating text based on user prompts. It reduces the time spent on brainstorming and drafting, allowing students to focus more on conceptualizing ideas and refining arguments.
2. **Access to Diverse Knowledge**: ChatGPT taps into a vast repository of information, drawing from various sources to provide comprehensive insights on a wide range of topics. This enables students to access diverse perspectives h findings into their papers.

Interacting with ChatGPT exposes students

Page 1 of 3 | 815 words | 153%

AI Writing

How much of this submission has been generated by AI?

86%

of qualifying text in this submission has been determined to be generated by AI.

Learn more

FAQs
View FAQs

Resources
Explore

Guides
View guides

Show disclaimer

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Result

# Results

**76%**

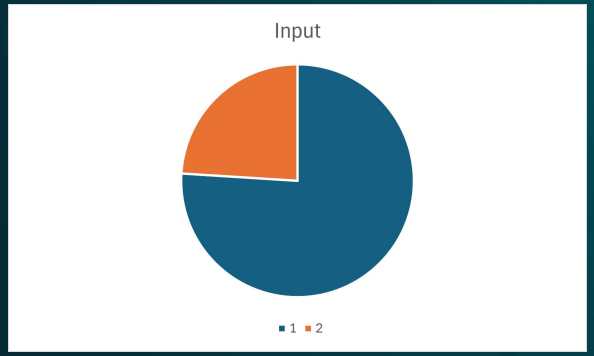
**623**
LLM words

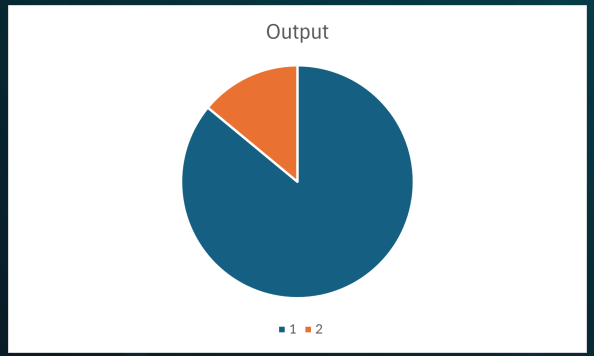
**25%**

**192**
Human words

**86%**

AI Similarities




The discrepancy of 10% not identified by the NLP.

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY

# Conclusion

***Can we answer the question of the study?***

***Can the unethical use of Gen AI Large Language models***

***be detected and conquered by Natural Language Processes?***

The study indicates that a GenAI LLM is a very powerful tool and can certainly be used by students and educators for finding, sorting and even summarizing text but should not be relied upon for writing the actual essay/ assignment.

The NLP is a programme that finds similarities with written work on the internet and produces a detailed report on what and from where text was copied and which and how much AI generated text was used.

.

# Conclusion

***Can we answer the question of the study?***

***Can the unethical use of Gen AI Large Language models***

***be detected and conquered by Natural Language Processes?***

This means that even though the NLP machines are brilliant and can present an astonishing result, human intervention is still very important for critical editing, proof reading and interpretation.

This study indicates the immense versatility of AI tools but in the same breath it also indicates the limitation of the machines and an important finding that human interaction is still needed.

***Machines may be able to "conquer" by detecting similar, plagiarised text but humans are needed to interpret the result and to make the final decision***.

# Conclusion

## Summary

NLP, Generative AI and LLMs, intertwine and diverge, shaping the path of modern AI.

**"Artificial intelligence is not a substitute for human intelligence; it is a tool to amplify human creativity and ingenuity."**

– Fei-Fei Li, Co-Director of the Stanford Institute for Human-Centered Artificial Intelligence

# Questions

**Thank you**

helga.hambrock@smu.ac.za
thandokazi.ikedinobi@smu.ac.za

SEFAKO MAKGATHO HEALTH SCIENCES UNIVERSITY