



# Revised Avenues of Assessment in Higher Education in the presence of AI Generative Contents

**Dr. Muhammad Iqbal**

Email: [miqbal@cct.ie](mailto:miqbal@cct.ie)

**ICT Faculty, CCT College Dublin, 30-34 Westmoreland St. Dublin 2, Ireland.**

- Introduction to Search engines and AI tools
- Recently launched AI tools
- Research Objectives
- Experiment
- Conceptual understanding of AI tools and behind technologies
- New or Revised methods of Assessments
- Comparison of Similarity results using zerochatGPT using Statistical plots
- Summary and Conclusions

# Introduction

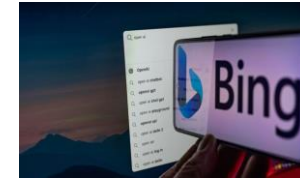
## Search Engines & AI chatbots



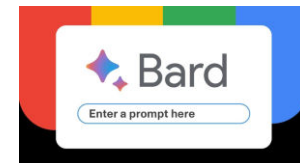
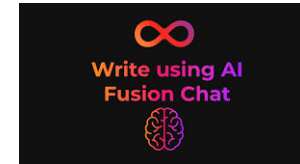
Users across  
different age groups



After  
2020



Before  
2020



# Introduction

## Various AI agents

- Generative Artificial Intelligence (AI) tools have been found to be effective and user-friendly in helping different kinds of users by generating a unique content based on a seed database in a variety of areas such as education, industry, and others.
- The use of AI generative tools will become more prevalent in the future.
- It will have a significant impact on the quality of learning in academia unless assessment standards are revised.
- In higher education, the continuous assessments are considered as a gauge throughout the semester to evaluate the learners' learning level, and if learner's response is based on AI generative contents.
- **Which tools are reliable to detect AI generated contents?**

chatGPT 3.5

Bing AI Chat

Fusion AI chat

Rix by  
Hashnode

Google bard

# Research Objectives

- Can ZerochatGPT be used as a Plagiarism detector for the identification of AI generated contents?
- To what extent different AI agents' answers differ from one another when asked similar questions?



- **Which methods can be utilised to evaluate the quality and authenticity of the assessments accurately?**



- **Experiment:** We designed 10 different questions from various categories.
- **Categories of Questions:** Blockchain, Big data, Deep Learning, Geography and Demographics, Earth Sciences, Social Progress and Well-being Studies, Chemistry, Anatomy and Physiology, Linguistics and Tourism Marketing
- We used 5 different AI tools to generate the response of these questions.
- ZerochatGPT used to determine the presence of AI contents in the generated responses.

# GPT Technology

## ChatGPT 3.5

- **chatGPT 3.5:** OpenAI was the company behind ChatGPT. ChatGPT is powered by GPT-3.5 at present, an acronym for "**Generative Pre-trained Transformer 3.5**".
- Transformer model used NN and this idea was introduced in a paper published in 2017.
- It has developed based on the use of natural language processing (NLP) and machine learning sequence-to-sequence models.
- Transformer Design is based on the model that focus on various segments of the input sequence while producing an output sequence, is its primary innovation.
- The transformer architecture is the backbone for many pre-trained language models, such as Google's BERT (Bidirectional Encoder Representations from Transformers) and OpenAI's GPT series, and it has shown to be quite effective for a variety of NLP tasks.

# GPT Technology

## Bing AI chat and Fusion AI chat

- **Bing AI Chat:**
  - Bing AI uses a combination of large language models (LLMs) and Microsoft Graph to provide an AI-powered chat experience. The LLMs are used to generate text-based on human inputs, making it ideal for chatbots and customer service interactions. Microsoft Graph is a unified API that provides access to data. By combining these two technologies, Bing Chat can provide a more personalized and efficient chat experience for its users.
- **Fusion AI chatbot:**
  - Fusion AI chat can "see" and "hear" like a human due to its AI speech and vision capabilities. With speech-to-text and text-to-speech capabilities, it can understand spoken language, analyse photos, and extract text using OCR technology. NLP powers Fusion AI chatbot and based on OpenAI's GPT (Generative Pre-trained Transformer) technology.



# GPT Technology

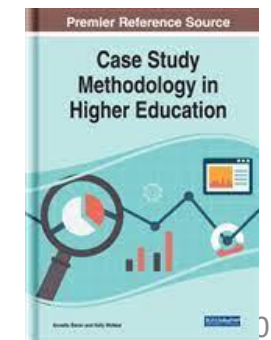
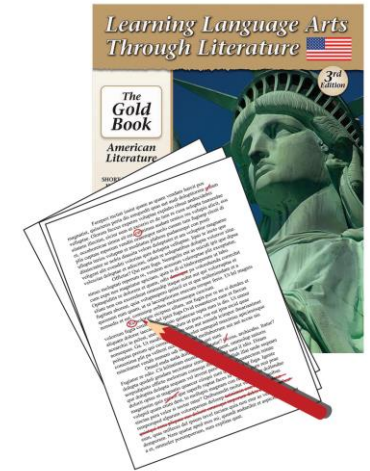
## RIX by hashnode and Google Bard

- **RIX by hashnode:**
- Rix is an AI-powered search engine built by Hashnode for developers. Rix leverages AI and machine learning techniques to provide relevant and accurate results, it does not specifically use OpenAI's GPT model.
- **Google Bard:**
- Google Bard is powered by Pathways Language Model 2 (PaLM 2). PaLM 2 is a 540-billion parameter language model, which means it has 540 billion parameters that can be tuned to learn and generate text. This makes PaLM 2 one of the largest and most powerful language models in the world.

# plagiarism Checker

## Technologies

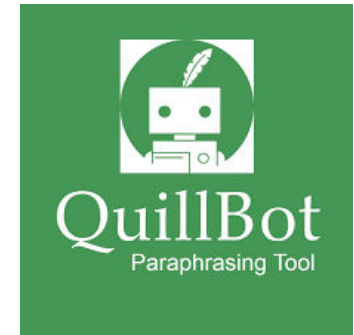
- A variety of evaluation methodologies, such as essay-based questions, case study-based activities, guided group work, and other strategies, and some of these need to be modernized in light of recent advances in AI content generation tools to avoid academic misconduct.
- **Major Technologies:** Turnitin, Grammarly, Copyscape, Plagscan, Quetext, Unicheck, DupliChecker, Viper and Copyleaks.
- In this paper, we explored the response obtained from free AI generative tools, such as **ChatGPT3.5, Microsoft Bing, Fusion AI Chatbot, Rix by hashnode** and **Google Bard** based on few queries and analysed output generated by them.
- **Can we consider ZerochatGPT as a reliable tool to detect the AI contents exactly?**



# New or Revised Methods

## Assessments

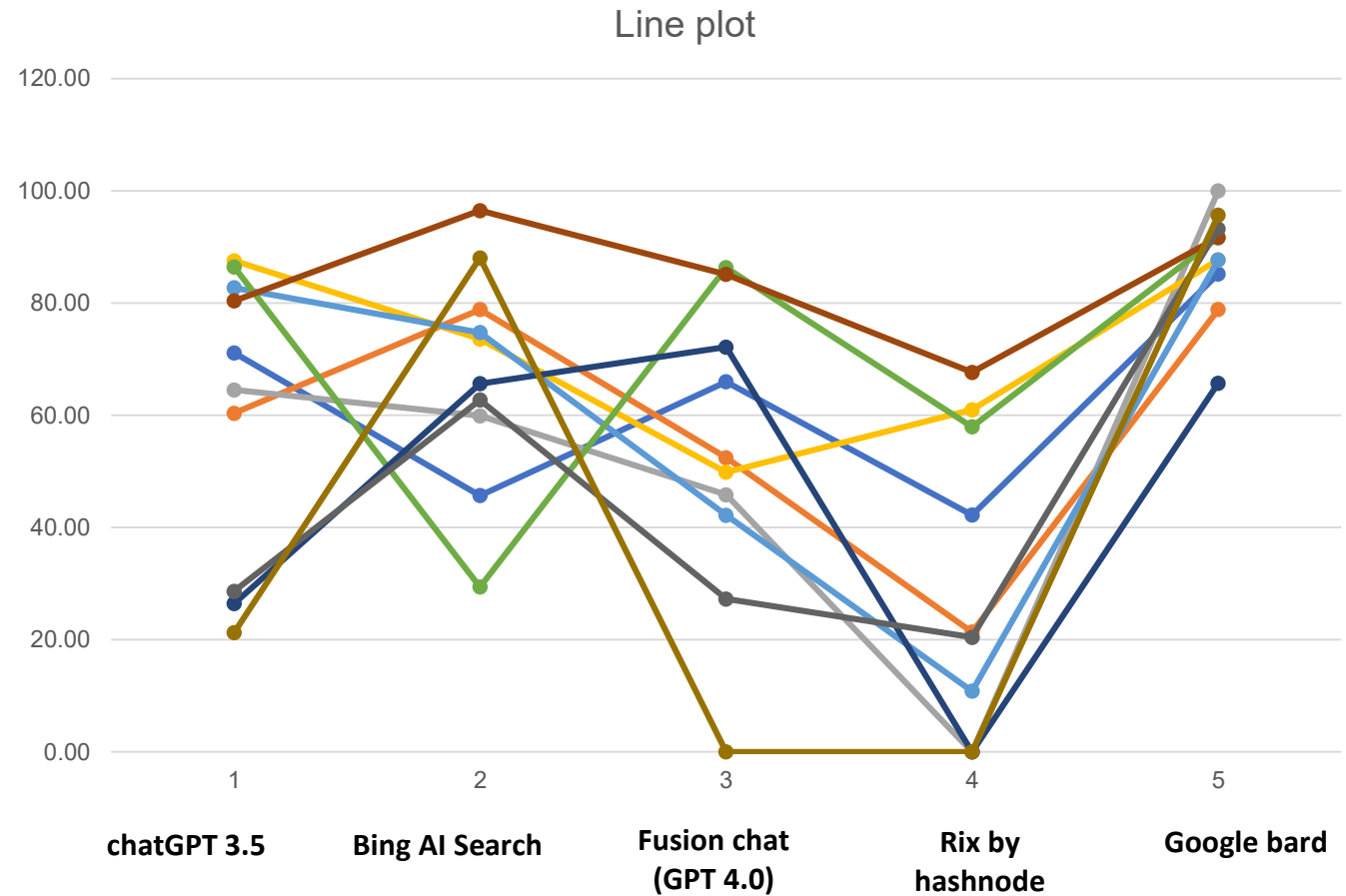
- **ZeroChatGPT** is employed to check the identification of AI generated contents from different AI chat tools and a distinct level of similarity is obtained for each AI tool.
- To determine the originality of academic work based on the use of a variety of AI generated contents along with tools like Quillbot is challenging in higher education.
- It has been discovered that the **ZerochatGPT** cannot correctly recognise the contents generated by chatGPT3.5, raising serious concerns about AI similarity detection technique.
- The analysis revealed that the use authentic assessment methods in addition viva voice/ presentation/ video recording along with explanation are essential to verify the correct level of learning in academic assessments in higher education.



# Comparison of Similarity

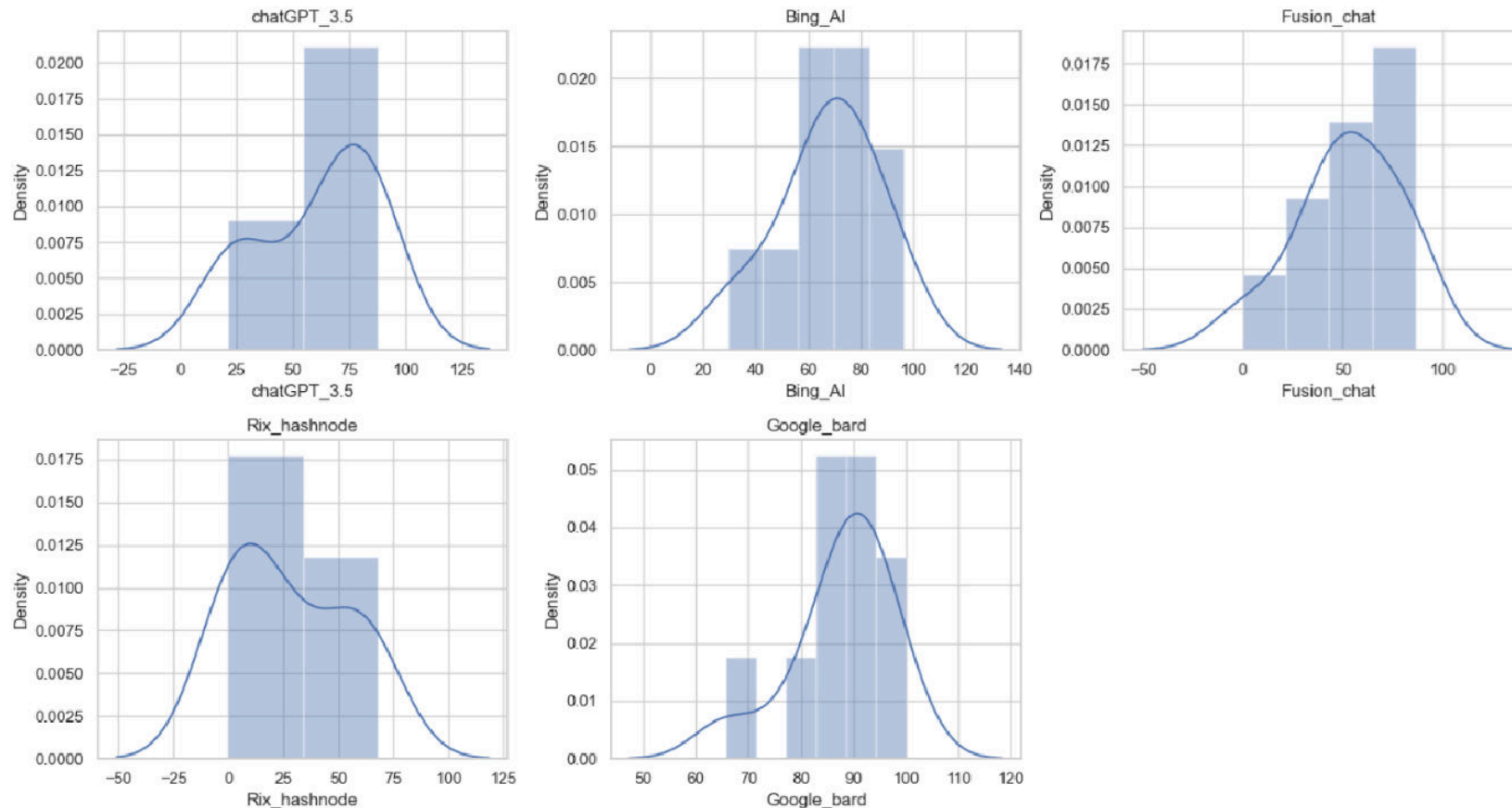
## zerochatGPT

- Text generated by different AI agents using various chat agents under various constraints.
- ZerochatGPT could not identify in some cases entirely that the text is generated by AI agent.
- The graph showed that the **Fusion chat** and **Rix by hashnode** AI generated text is difficult to identify in some cases and showed 0% similarity.



# Distribution plots

## zerochatGPT

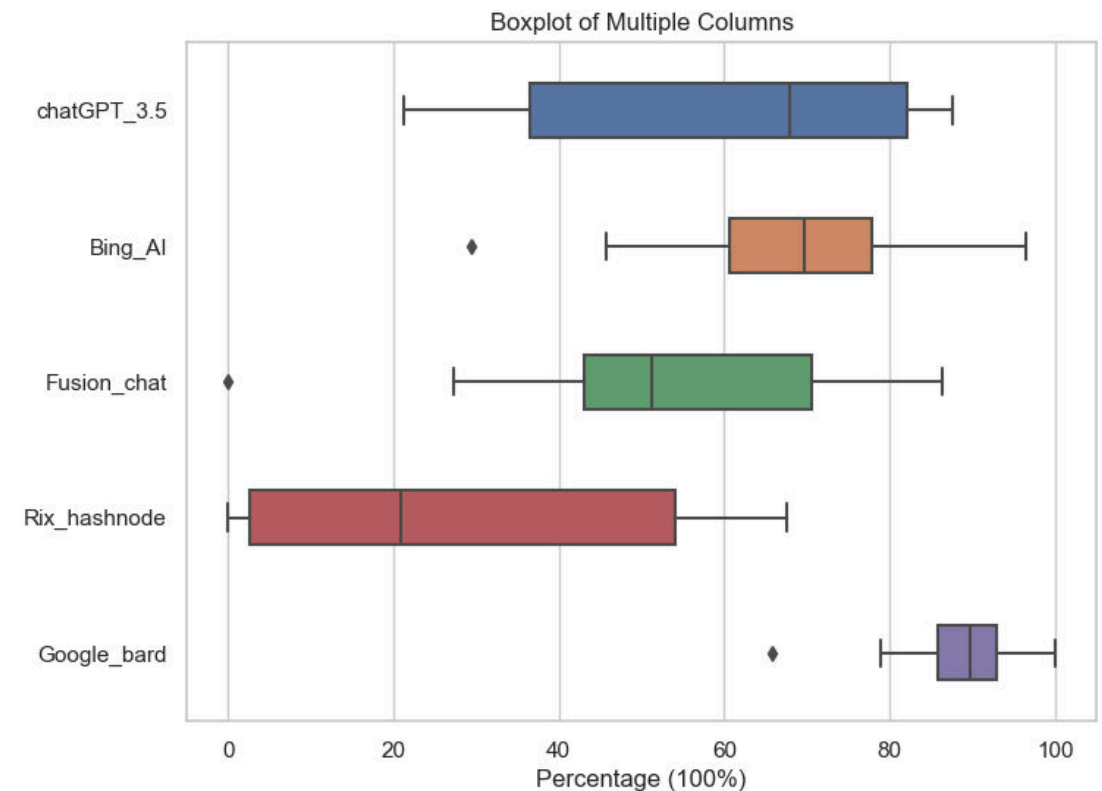


- The distribution plots clearly showed that Rix by hashnode and Fusion AI started from 0% detection by zerochatGPT.

# Box plot and Outliers

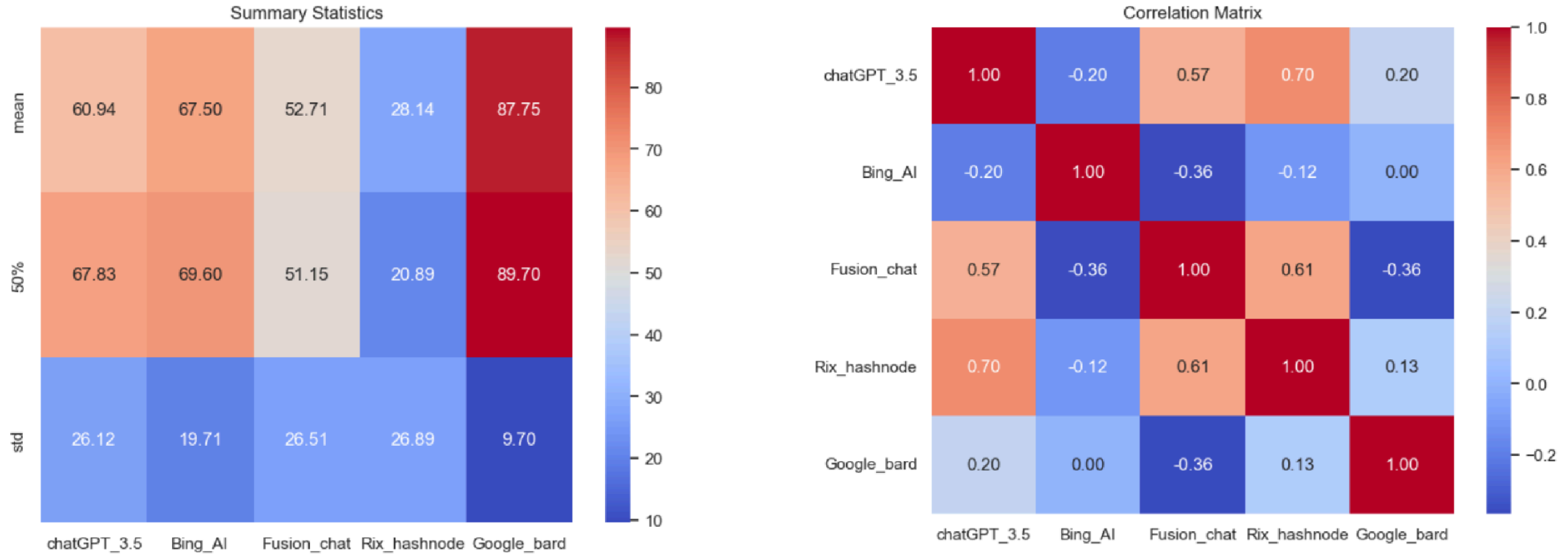
## zerochatGPT

- It is clear from the boxplot that the outliers are present in Bing AI, Fusion AI chat and Google bad and there are not any outliers in the case of Rix by hashnode.
- The values generated by Rix hashnode is a challenge to identify for the zerochatGPT.
- **Challenge:** The evaluation of the response/text quality generated by AI chatbots, which was based on 10 distinct questions, requires further research and comprehension and should be examined by subject matter experts in the relevant domains.



# Summary and Correlation Statistics

## zerochatGPT



- The mean value of Rix hashnode is significantly lower than other AI chat agents. A standard deviation of 26.89 indicates that, on average, individual data points deviate from the mean by approximately 26.89 units in the case of Rix by hashnode.
- Pearson Correlation for Rix by hashnode is the lowest as compared to Google bard and highest with chatGPT3.5.

# Summary and Conclusions

- We identified the limitations for the usage of zerochatGPT as a similarity detection tool.
- It is not reliable to check the student's assessment work based on zerochatGPT for the detection of plagiarism.
- We must use the oral methods, such Viva voice, presentation and video recordings to verify the authenticity of assessment.
- We provided the statistics of the results generated by 5 different AI agents based on 10 distinct questions from science and arts fields.
- We can say that this area needs to be explored further and the quality of response generated by AI chatbots should be analysed by the subject matter experts.
- A further exploration and detailed analysis is required to understand the quality of the text generated by various AI chat tools.



# Resources/ References

- Muhammad Iqbal, Presentation, Engaging Students with Authentic Assessment, CCT Teachmeet (2022), CCT College Dublin.
- GPT 3.5, Web: <https://chat.openai.com>
- Bing AI chat with GPT4, Web: Bing Chat with GPT-4
- AI Assistant, Fusion Chat, <https://fusionchat.ai>
- RIX by hashnode, Web: <https://hashnode.com>
- Google Bard, Web: <https://bard.google.com>
- Will ChatGPT get you caught? Rethinking of Plagiarism Detection, Mohammad Khalil and Erkan Er, Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway, Middle East Technical University, Turkey.
- Some images are used from Google Search repository to enhance the level of learning for educational purpose.
- Python code developed for Statistical analysis.