ISSN (Print) : 0972-2467
ISSN (Online) : 0976-2477
Journal of Information and Knowledge, Vol 60(3), June 2023, p.175-183
DOI: 10.17821/srels/2023/v60i3/171028

Scholarly Communication and Machine-Generated Text: Is it Finally AI vs AI in Plagiarism Detection?

Patit Paban Santra[1]* and Debasis Majhi[2]

[1]Assistant Librarian, DSMS College, Durgapur – 713206, Durgapur, West Bengal, India; ppsantra95@gmail.com
[2]Research Scholar, Banaras Hindu University, Varanasi – 221002, Uttar Pradesh, India; debasismajhidlis@gmail.com

Abstract

This study utilizes GPT (Generative Pre-Trained Transformer) language model-based AI writing tools to create a set of 80 academic writing samples based on the eight themes of the experiential sessions of the LTC 2023. These samples, each between 2000 and 2500 words long, are then analyzed using both conventional plagiarism detection tools and selected AI detection tools. The study finds that traditional syntactic similarity-based anti-plagiarism tools struggle to detect AI-generated text due to the differences in syntax and structure between machine-generated and human-written text. However, the researchers discovered that AI detector tools can be used to catch AI-generated content based on specific characteristics that are typical of machine-generated text. The paper concludes by posing the question of whether we are entering an era in which AI detectors will be used to prevent AI-generated content from entering the scholarly communication process. This research sheds light on the challenges associated with AI-generated content in the academic research literature and offers a potential solution for detecting and preventing plagiarism in this context.

Keywords: AI (Artificial Intelligence), GPT (Generative Pre-Training Transformer), Machine Learning, ChatGPT, Natural Language Processing (NLP), OpenAI, Plagiarism

1. Introduction

Similarity-based anti-plagiarism tools are based on two concepts: one is semantic similarity which refers to the degree to which two pieces of text, or two concepts, have a similar meaning (Chowdhury and Bhattacharyya, 2018). It is a measure of the closeness of the meanings of two words or concepts based on their relationships in a given context. On the other hand, syntactic similarity refers to the degree to which two pieces of text are similar in structure or arrangement of words (Oya, 2020). This can be evaluated by comparing the syntax or grammatical structure of the two texts. AI-generated content is produced by machines using artificial intelligence against a suitable prompt or instruction. AI generators require some given text as an initial, like a description, prompt, or parameters. Depending on the initial text, an AI generator can create a paragraph or large section of text. Nowadays, for time-saving, efficiency, and perfect construction of text, AI generator tools are used widely in blog posts, ad copy, product descriptions, marketing copy, and articles. The ELIZA language model, created in the 1960s at MIT, US, by Joseph Weizenbaum, is one of the earliest instances of a language model for computer-generated writing (Weizenbaum, 1966). Presently, the domain of AI writing is dominated by three major language models:

- GPT (Generative Pre-training Transformer): Developed by OpenAI in 2018, GPT is a transformer-based language model that uses unsupervised learning to generate human-like text based on a given prompt (Topal *et al.*, 2021). GPT has been widely used for a variety of tasks, including language translation, question answering, and text generation.

*Author for correspondence

- BERT (Bidirectional Encoder Representations from Transformers): Developed by Google in 2018, BERT is a transformer-based language model that uses unsupervised learning to generate high-quality text representations that can be used for a variety of natural language processing tasks (Birunda and Devi, 2021). BERT has been widely used for tasks such as language translation, question answering, and text classification.
- RoBERTa (Robustly Optimized BERT Approach): Developed by Facebook in 2020, RoBERTa is a variant of BERT that was designed to improve upon the original model by using more data and more computing resources during training. RoBERTa has been shown to perform well on a variety of natural language processing tasks, including language translation, question answering, and text classification (Cortiz, 2022).

The GPT language model requires a special mention here not only for its technical capabilities but also for its openness. While the GPT model itself is not open access or open standard, the training data and code used to train the model have been made publicly available by OpenAI, and the model can be accessed and used by anyone through the OpenAI API, and thereby making it an influential tool in the field of AI/ML based text generation (Chan, 2023). The GPT model has so far four major milestones, and these are:

- GPT-1: GPT-1 uses a transformer architecture, which is a type of neural network architecture that was introduced in 2017 by OpenAI and has since become widely used in natural language processing tasks.
- GPT-2: GPT-2 is a transformer-based language model that was released by OpenAI in 2019. GPT-2 is notable for its large size, with over 1.5 billion parameters, making it one of the largest language models available at the time of its release.
- GPT-3: It is an improvement over GPT-2 (released in 2020), with a larger size (over 175 billion parameters) and more advanced training techniques that allow it to generate higher-quality text and perform a wider range of natural language processing tasks. GPT-3 can be fine-tuned for tasks such as language translation, question answering, and text generation and has been used in a variety of applications, including chatbots, content generation, and language translation.
- GPT-3.5: GPT-3.5 is a variant of GPT-3 that was released by OpenAI in 2021. It is similar to GPT-3 in terms of its size and capabilities but has been further optimized and fine-tuned to improve performance on a variety of natural language processing tasks.
- GPT-4: It is the successor to GPT-3 and now it also the latest version of GPT model. It was officially launched on 14th March 2023 by OpenAI. It can process up to 25,000 words, almost eight times as many as GPT-3. This latest version also processes images and handles more distinction instructions than GPT-3.5.

ChatGPT, the newest entry in the series (released in November 2022), is a chatbot that uses the GPT-3 language model (GPT-3.5 to be more exact, though not clearly stated in the release document) developed by OpenAI to generate responses to user input. It is an example of how the GPT-3 language model can be used to build interactive chatbots that can engage in conversation with users and generate human-like responses.

2. Related Literature

Computer-generated content first focused on the visual arts and music fields in the late 1950s (Boden and Edmonds, 2010), and the first machine-generated book was published by Springer Nature in collaboration with researchers from Goethe University, Frankfurt, Germany in 2009 (Writer, 2019). But at that time, it was possible to easily detect machine-generated or computer-generated text and researchers could easily differentiate between computer-generated text, and human-created text (Pataranutaporn *et al.*, 2021). But the sliding problem arrived when Natural Language Processing (NLP) based large language model was developed for creating computer-generated content. Petroni *et al.* analyzed that the pre-trained language model not only learned from the store dataset but also stored the knowledge for future use or similar use in the future (Petroni *et al.*, 2019). In the 2020s, the Generative Pre-trained Transformer 3 (GPT-3) based model was developed by an Open AI organization for generating the AI text (Brown *et al.*, 2020). Using the GPT-3 based model such as ChatGPT can generate N number of machine-generated text just by putting some topic, idea, phases, or keywords, and it is humanly understandable, meaningful, and harmless text content (Crothers *et al.*, 2023). The most important thing in scientific communication is the scientific paper, and these papers reflect the researcher's quality, researcher ability, and working knowledge in his or her fields and from where the peoples and other researchers gain knowledge about things. But AI writing tools, such as ChatGPT are increasing their presence in scientific writing fields and have raised concerns among the scientific community due to the ethical questions it raises. The first computer-generated paper was identified in 2005 that was generated through SCIgen and published by reputed academic publishers such as Springer and IEEE (Van Noorden, 2014)**.** In 2013, Labbé & Labbé developed a model based on the text mining process to deceive the machine-generated fake papers. Another process was identified by Oberreuter and Velásquez (2013) based on the writing style of text and calculating the frequency of words to detect the text that was not plagiarized in plagiarism-detected tools. Recently pre-trained language model like GPT-3 can even write a paper about itself with the title 'Can GPT-3 write an academic paper on itself, with minimal human input? (Transformer *et al.*, 2022). Also, some researchers use ChatGPT as co-author of their papers (A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education; Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse?) and published in reputed Journal like *Nature*, Springer. (King and chatGPT, 2023; O'Connor and ChatGPT, 2023). In this situation, every educational and research organization is worried about machine-generated text or ChatGPT-generated content. The existence of computer-generated content in the academic community has been observed since the early 2000s, and it was initially easy for researchers and reviewers to identify such content. However, with the emergence of GPT-3 based generated content, traditional plagiarism detection tools are unable to detect them, raising concerns about the potential publication of such content in reputed journals. With the issue of plagiarism detection for machine-generated content in mind, this paper aims to explore the gap and solution to the problem.

3. Objectives

In view of the possible disruption that powerful AI writing tools like ChatGPT can create in the scholarly communication process, the objectives of this study are as follows:

- To examine the ability and limitations of typical plagiarism detectors to identify machine-generated scholarly text;
- To identify tools that can detect machine-generated scholarly text; and
- To explore the applications of these tools in identifying machine-generated text.

4. Methodology

Artificial Intelligence (AI) can generate various types of content much faster than humans, making it difficult to distinguish between human and AI-generated content. To address this challenge, an exploratory investigation was conducted using AI detector tools and AI-generated text. A three-step methodology was adopted: the first step focused on creating AI/machine-generated text, while the second and third steps were focused on detecting plagiarism in the machine-generated text.

First: We created a set of 80 AI-generated content using ChatGPT with all typical sections and styles in the form of scholarly articles based on eight (8) themes of experiential sessions of LTC 2023 containing an introduction, objectives, methodology, conclusion, etc. Each theme had ten (10) textual objects related to different facets and sub-facets of the given theme (2000-2500 words).

Second: We examined these AI-generated content one by one with the conventional plagiarism detection tools (Turnitin, iThenticate, Ouriginal-Urkund, and Duplichecker) to find out whether these conventional plagiarism tools can detect the AI-generated text or not.

Third: We applied the same principle with a few selected AI detector tools (GPT-2, Content at Scale, Writer.com, etc.) that are based on GPT-3 model-based and examined if the same text generated by ChatGPT can be detected or not. We additionally went into detail on the advanced Artificial intelligence detector tools, along with how they work and how they may detect text corpora that were generated by AI.

Table 1. AI Content dictation of abstract fields of SRELS Journal of Information Management

Sample data (n = 7) abstract-wise (Vol. No. 59, Issue No. 06)	AI-Content dictation through GPT-2 tool	
	Real (%)	**Fake (%)**
(Roy and Mukhopadhyay, 2022)	99.86	0.14
(Pal and Mukhopadhyay, 2022)	99.97	0.03
(Maity and Dutta, 2022)	99.98	0.02
(Oladokun *et al.*, 2022)	99.98	0.02
(Wani and Bhat, 2022)	99.98	0.02
(Parmar and Nagi, 2022)	83.67	16.33
(Das *et al.*, 2022)	99.98	0.02

Also, for evaluating the AI detector tools, they were applied to all the articles (n = 7) published in the December 2022 issue of the *SRELS Journal of Information Management* and check how AI tools detected the human-generated text (Table 1).

5. Results

It was found that most of the comprehensive similarity detection software for research communication, the Ouriginal-Urkund, the Turnitin, the iThenticate, and the Duplichecker have failed repeatedly to detect scholarly text produced by ChatGPT writing tool (Figure 1). Traditional syntactic similarity-based anti-plagiarism tools may indeed have difficulty detecting AI-generated text, as the text produced by AI models can often be quite different from the human-written text in terms of syntax and structure. This is because AI models are not bound by the same grammatical and syntactic rules as humans and may generate text that is more akin to machine-generated code than to human-written language.

But the same machine-generated content can be detected by a freely available AI detector (Figure 2). AI detectors are algorithms or systems that are designed to identify and classify text or other content that has been generated by an Artificial Intelligence (AI) system, as opposed to content produced by a human. Some principles and methods are used in the design and operation of AI detectors:

- Natural Language Processing (NLP): AI detectors often rely on techniques from the field of NLP to analyze the structure, syntax, and semantics

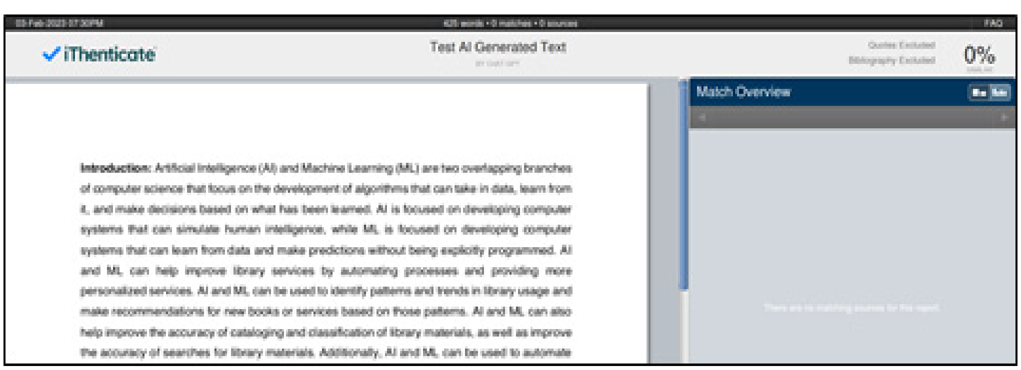


Figure 1. Similarity (0%) report for AI-generated scholarly text in iThenticate.

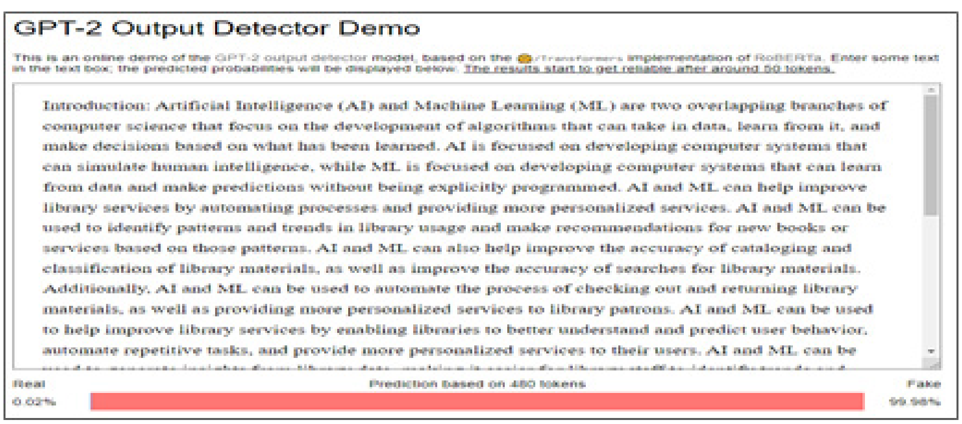
of written or spoken language. By analyzing the patterns and characteristics of language, AI detectors can identify patterns or features that are indicative of machine-generated text.

- Machine learning: Many AI detectors use machine learning algorithms to learn and adapt over time, improving their ability to accurately identify machine-generated text.
- Statistical analysis: This approach uses statistical analysis to identify patterns in the text that are indicative of machine generation. For example, machine-generated text may have a different distribution of word lengths or frequencies of certain words compared to human-written text.
- Style analysis: Another approach is to use style analysis to identify characteristics of the writing style that are indicative of machine generation. For example, machine-generated text may use a more formal or repetitive style compared to human-written text.
- Content analysis: AI detectors may also analyze the content of the text or other media to identify patterns or features that are indicative of machine-generated content. This can involve analyzing

Table 2. Similarity score comparison between Plagiarism tools and AI-detector tools

Anti Plagiarism Tools	Score (%)		AI-detector Tools	Score (%)	
	Similarity	Unique		Real (Human Content)	Fake (AI Content)
Ouriginal	0	100	GPT-2	0.02	99.98
Turnitin	4	96	Content at Scale	0	100
iThenticate	0	100	Writer.com	1	99
Dupli Checker	0	100	Sapling.ai	0	100

#Unique: Content free from plagiarism



Source: https://openai-openai-detector.hf.space/

Figure 2. Real (0.02%) vs Fake (99.98%) report for AI-generated scholarly text in GPT detector demo.

the topic, tone, or style of the content, as well as examining more technical aspects of the text, such as the use of specific words or phrases.

- Human evaluation: In some cases, AI detectors may be evaluated and refined based on feedback from human evaluators, who can provide insights into the performance of the AI detector and help identify areas for improvement.

In short, AI detectors rely on a combination of these and other principles and methods to accurately identify and classify machine-generated text.

It is quite interesting to note that highly paid anti-plagiarism tools fail to detect machine-generated text, but freely available AI detector tools can identify it easily. There are many approaches that can be used by these tools to detect AI-generated text; potential methods are statistical analysis, style analysis, vocabulary, grammar, etc. An AI detector tool is specifically designed to identify whether a piece of text has been generated by an AI system or a human being. These tools can analyze various features of the text, such as the language style, sentence structure, and patterns of repetition, to determine whether the text was created by a machine learning algorithm or a human writer. On the other hand, a plagiarism detector tool is designed to compare a given piece of text with other sources to identify similarities and potential instances of plagiarism. While these tools can be useful for identifying text that has been copied from other sources, they are not specifically designed to detect AI-generated text. Therefore, if one is looking to detect whether a piece of text has been generated by an AI system, an AI detector

Table 3. Theme-wise overall results (n = 80)

Themes of LTC_2023	**AI-Generated Text corpus (2000-2500 words)**	**Similarity Score (%) [Selected Anti plagiarism tools]**		**AI-Content (Fake) Score (%) [Selected AI-Detector tools]**	
		Highest	**Lowest**	**Highest**	**Lowest**
Library Service Platform	10	2	0	100	99.98
Data Deluge	10	3	0	100	99
E-Resource Management	10	1	0	100	100
Innovative Library Services and Promotions	10	4	0	100	100
Web-enabled Library Systems	10	3	0	100	99
AI/ML applications	10	3	0	100	99.97
Open Access (OA) and Libraries	10	2	0	100	99.96
E-learning and Libraries	10	4	0	100	99.98
Total Sample Size	**80**				

tool would be more appropriate than a plagiarism detector tool.

Based on Table 2 results, it is possible to draw a comparison between plagiarism detector tools and AI-detector tools. According to methodology, a few AI-generated texts were checked through anti-plagiarism tools, and the same texts with AI-detector tools. The performance of plagiarism tools is very poor, showing less than 1 per cent plagiarism in most cases. AI detector tools are showing fake (AI content) scores of more than 99 per cent only less than 1 per cent is real (human content).

6. Overall Result (n = 80)

Each of the AI-generated text (n = 80) was checked against plagiarism using four traditional anti-plagiarism software. The similarity score varied between 0% and 4% in all cases. On the other hand, same text corpora were checked against fake (AI content) score using four AI detectors. The fake (AI content) score varied between 99% and 100%, which means most of the content are generated by AI writing tool (ChatGPT), human-generated content is less (Table 3).

Table 3 also shows the theme (LTC 2023) wise results for a total of 80 AI-generated texts against plagiarism and AI detection. Based on 8 themes of LTC 2023, total of 80 AI-generated text were prepared using ChatGPT. There were 8 themes and a set of 10 AI-generated texts was generated from each theme. Table 3 also shows the theme-wise highest and lowest similarity scores and fake scores. The highest similarity score is found between 0% and 4% in the 4^{th} and 8^{th} themes, 'Innovative Library Services and Promotions' and 'E-learning and Libraries', and the lowest similarity score is found between 0% and 1% in the 3rd theme, 'E-Resource Management'. The average similarity score is 2.75 among all 8 themes. In respect of the similarity report of all samples, we may say that the performances of plagiarism software is poor, and these tools fail to detect AI-generated text. The lowest AI content (fake) score is found to be between 99% and 100% in the 2nd and 5th themes, namely 'Data Deluge' and 'Web-enabled Library Systems'. So, the fake content score is above 99%, and the real one is below 1%. According to table data, theme numbers 3rd and 4th are in the top position with 100% fake score and 0% human content score on both sides (Highest and Lowest). It is proved that all contents are generated through GPT based AI writing tool (ChatGPT).

Finally, it was observed that human text and AI-generated text are almost indistinguishable for humans. Nonetheless, AI detection tools can identify specific techniques, structures, and machine learning algorithms used in AI-generated text. Additionally, this research identifies ethical measures that can assist in maintaining academic integrity and research ethics.

7. Study Limitations

The study has some limitations. Firstly, it only examines a single AI writing model, ChatGPT, while there are many other AI models/tools available on the Internet. Secondly, the accuracy of the results depends on the precision of four plagiarism detection software tools, namely Ouriginal-Urkund, Turnitin, iThenticate, and Dupli Checker. Thirdly, the sample size of 80 AI-generated texts may not provide a comprehensive perspective. A larger sample size of over 500 may be necessary, and further research may be required to generalize the findings of this study.

8. Conclusion

This study raises the question of whether machine-generated scholarly text tools, based on the GPT-3 model, such as ChatGPT, can disrupt future research and scholarly communication processes. These tools have the potential to be disruptive technologies as the GPT-3 model is a highly advanced language model that can generate text that is nearly indistinguishable from human writing. Consequently, tools like ChatGPT could be utilized to produce scholarly articles, conference papers, and other types of written works. Moreover, the research findings reveal that traditional plagiarism detection tools, including the most powerful ones, cannot detect machine-generated text. Thus, there is a growing need to leverage AI for plagiarism detection, which may eventually result in 'AI vs AI' in the fight against plagiarism in scholarly communication.

9. References

Birunda, S. S. and Devi, R. K. (2021). A review on word embedding techniques for text classification. In J. S. Raj, A. M. Iliyasu, R. Bestak, and Z. A. Baig (Eds.), Innovative Data Communication Technologies and Application, p. 267-281. https://doi.org/10.1007/978-981-15-9651-3_23

Boden, M. A. and Edmonds, E. A. (2009). What is generative art? Digital Creativity, *20*(1-2), 21-46. https://doi.org/10.1080/14626260902867915

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners, Available at: https://arxiv.org/abs/2005.14165.

Chan, A. (2023). GPT-3 and InstructGPT: Technological dystopianism, utopianism, and 'Contextual' perspectives in AI ethics and industry. AI and Ethics, *3*(1), 53-64. https://doi.org/10.1007/s43681-022-00148-6

Chowdhury, H. A. and Bhattacharyya, D. K. (2018). Plagiarism: Taxonomy, tools and detection techniques.

Cortiz, D. (2022). Exploring transformers models for emotion recognition: A comparision of BERT, DistilBERT, RoBERTa, XLNET and ELECTRA. Proceedings of the 2022 3rd International Conference on Control, Robotics and Intelligent System, 230-234. https://doi.org/10.1145/3562007.3562051

Crothers, E., Japkowicz, N., and Viktor, H. (2023). Machine generated text: A comprehensive survey of threat models and detection methods. Available at: https://arxiv.org/abs/2210.07321

Das, A., Mandal, N., Rath, D. S. and Das, S. (2022). Trendline of open access publication by Indian Institute of Technology (IITs) researchers in India. SRELS Journal of Information Management, 399-409. https://doi.org/10.17821/srels/2022/v59i6/168621

King, M. R. and chatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. Cellular and Molecular Bioengineering, *16*(1), 1-2. https://doi.org/10.1007/s12195-022-00754-8

Labbé, C. and Labbé, D. (2013). Duplicate and fake publications in the scientific literature: How many SCIgen papers in computer science? Scientometrics, *94*(1), 379-396. https://doi.org/10.1007/s11192-012-0781-y

Maity, D. and Dutta, B. (2022). Identifying the core and allied disciplines involved in the growth of virology: A linguistic analysis. SRELS Journal of Information Management, 363-371. https://doi.org/10.17821/srels/2022/v59i6/170750

Oberreuter, G. and Velásquez, J. D. (2013). Text mining applied to plagiarism detection: The use of words for detecting deviations in the writing style-ScienceDirect. Expert Systems with Applications, *40*(9), 3756-3763. https://doi.org/10.1016/j.eswa.2012.12.082

O'Connor, S. and ChatGPT. (2023). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? Nurse Education in Practice, *66*. https://doi.org/10.1016/j.nepr.2022.103537

Oladokun, B. D., Seidu, A. E., Ogunbiyi, J. O., Aboyade, W. A., Yemi-Peters, O. E. and Elai, M. A. (2022). Utilization of Information and Communication Technologies (ICTs) for managing students' academic records in Nigerian Schools. SRELS Journal of Information Management, 373-381. https://doi.org/10.17821/srels/2022/v59i6/168449

Oya, M. (2020). Syntactic similarity of the sentences in a multi-lingual parallel corpus based on the Euclidean distance of their dependency trees. Proceedings of the 34th Pacific Asia Conference on Language, Information, and Computation, 225-233.

Pal, A. and Mukhopadhyay, P. (2022). Fetching automatic authority data in ILS from Wikidata via OpenRefine. SRELS Journal of Information Management, 353-362. https://doi.org/10.17821/srels/2022/v59i6/170677

Parmar, R. D. and Nagi, P. K. (2022). Institutional knowledge repositories: Re-contextualization for accreditation and quality management. SRELS Journal of Information Management, 383-390. https://doi.org/10.17821/srels/2022/v59i6/170796

Pataranutaporn, P., Danry, V., Leong, J., Punpongsanon, P., Novy, D., Maes, P. and Sra, M. (2021). AI-generated characters for supporting personalized learning and well-being. Nature Machine Intelligence, *3*(12). https://doi.org/10.1038/s42256-021-00417-9

Petroni, F., Rocktäschel, T., Riedel, S., Lewis, P., Bakhtin, A., Wu, Y. and Miller, A. (2019). Language models as knowledge bases? Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2463-2473. https://doi.org/10.18653/v1/D19-1250

Roy, B. K., and Mukhopadhyay, P. (2022). Digital access brokers: Clustering and comparison (Part II - from Summarization to Citation Map). SRELS Journal of Information Management, 337-351. https://doi.org/10.17821/srels/2022/v59i6/170786

Topal, M. O., Bas, A. and van Heerden, I. (2021). Exploring transformers in natural language generation: GPT, BERT, and XLNet. Available at: https://arxiv.org/abs/2102.08036

Transformer, G. G. P., Thunström, A. O. and Steingrimsson, S. (2022). Can GPT-3 write an academic paper on itself, with minimal human input?

van Noorden, R. (2014). Publishers withdraw more than 120 gibberish papers. Nature. https://doi.org/10.1038/nature.2014.14763

Wani, Z. A. and Bhat, A. (2022). Figshare: A one-stop shop for research data management with diverse features and services. SRELS Journal of Information Management, 391-397. https://doi.org/10.17821/srels/2022/v59i6/170789

Weizenbaum, J. (1966). ELIZA-a computer program for the study of natural language communication between man and machine. Communications of the ACM, 36-45. https://doi.org/10.1145/365153.365168

Writer, B. (2019). Lithium-ion batteries: A machine-generated summary of current research. Springer International Publishing. https://doi.org/10.1007/978-3-030-16800-1