Is ChatGPT a Reliable Source of Transportation Equity Information for Scientific Writing?

Boniphace Kutela, Ph.D., P.E.[1]; Shoujia Li[2]; Subasish Das, Ph.D.[3]; and Jinli Liu[4]

[1]Roadway Safety Program, Texas A&M Transportation Institute, Houston, TX.
Email: b-kutela@tti.tamu.edu
[2]Texas A&M Univ., College Station, TX. Email: shoujiali@tamu.edu
[3]Texas State Univ., San Marcos, TX. Email: subasish@txstate.edu
[4]Texas State Univ., San Marcos, TX. Email: jinli.liu@txstate.edu

ABSTRACT

Transportation equity is an interdisciplinary agenda that requires both transportation and social inputs. Traditionally, transportation equity information is sourced from public libraries, conferences, television, and social media, among others. Artificial intelligence (AI) tools, including advanced language models such as ChatGPT, are becoming favorite information sources. However, their credibility has not been well explored. This study explored the content and usefulness of ChatGPT-generated information related to transportation equity. It utilized 152 papers retrieved through the Web of Science (WoS) repository. The prompt was crafted for ChatGPT to provide an abstract given the paper's title. The ChatGPT's abstracts were then compared to human-written abstracts using statistical tools and unsupervised text mining. The results indicate a weak similarity between ChatGPT and human-written abstracts. On average, the human-written and ChatGPT-generated abstracts were about 58% similar, with a maximum and minimum of 97% and 1.4%, respectively. The keywords from the abstracts of papers with over the mean similarity score were more likely to be similar, whereas those below the average score were less likely to be similar. Themes with high similarity scores include access, public transit, and policy, among others. Contrarily, the findings from collocated keywords were inconclusive. The study findings suggest that ChatGPT has the potential to be a source of transportation equity information. However, currently, a great amount of attention is needed before a user can utilize materials from ChatGPT.

Keywords: Transportation equity; Artificial intelligence; ChatGPT

INTRODUCTION

The use of artificial intelligence (AI) is radically altering how individuals live their lives and carry out the many activities they do daily. People's ways of living and working have been revolutionized because of the widespread adoption of cutting-edge technologies such as smartphones and smart watches. In addition, the use of voice command devices like SIRI and Alexa has transformed the day-to-day activities that people participate in.

In November 2022, OpenAI released ChatGPT, an advanced language model, that interacts with users by providing a set of written directives and produces the written text according to the instruction given (Noever & Ciolino, 2022). This tool has gained the attention of experts in different fields, such as academicians, economists, social scientists, engineers, and computer scientists (ChatGPT & Perlman, 2022; Gao et al., 2022). Most of the concerns surrounding this tool have been about whether ChatGPT will replace some human-generated activities, such as writing codes/algorithms, preparing poems, movie transcripts, etc. (Qadir, 2022). Some scholars even argue that certain jobs will be replaced by ChatGPT (Qadir, 2022), whereas others disagree by indicating that the tool is not capable of taking over most human-generated jobs and thus will have minimal impact (Aydın & Karaarslan, 2022; Frye, 2022). However, ChatGPT is based on the advanced language model, which uses reinforcement learning, so it is expected to get better when new observations are included in retraining the model.

There has been discussion from researchers and the public about the replacement of Google with ChatGPT (Bindra, 2023; Greene, 2023). Google searches have been used by various researchers as the starting point for searching for various information. A key question among researchers has been whether the future of Google searches is on the brink of collapsing. Due to its ability to create informative texts given a certain prompt, researchers and the public have been using ChatGPT as a source of information. Users have prompted topics related to politics, social science, comedy, business, art, healthcare, games, coding, computer science, and transportation, among others (Kim & Lee, 2023; Yalalov, 2022). In the transportation field, public transportation accessibility and affordability have been among a few areas where users have been searching for information using ChatGPT (Mobility-Innovators, 2023). Public transportation accessibility and affordability form a part of transportation equity, which is among the hot topics both in the United States and globally.

Transportation equity is among the emerging topics of interest in the transportation field. The concept of transportation equity pertains to guaranteeing fair access to transportation alternatives for all individuals, regardless of their socioeconomic position or other individual circumstances (di Ciommo & Shiftan, 2017). With the global trend of urbanization leading to a surge in population and associated problems such as traffic gridlock and environmental pollution, transportation equity is gaining prominence as a pressing issue in various regions. Transportation equity is crucial as it has a pivotal role in advancing social and economic inclusivity (di Ciommo & Shiftan, 2017; Martens et al., 2019). Access to transportation is critical for individuals to access essential services like education, healthcare, and employment opportunities. Inadequate transportation infrastructure or unaffordable transportation options can impede people's ability to access such opportunities(Martens et al., 2019). By enhancing transportation choices and ensuring equal access, transportation equity can contribute to reducing poverty, improving social mobility, and boosting economic development.

Considering the importance of transportation equity, various outlets have been used as sources of information for this topic. Such sources include the National Association of City Transportation Officials (NACTO), the National Center for Mobility Management (NCMM), the American Public Transportation Association (APTA), and the U.S. Department of Transportation (USDOT). Other sources of information for transportation equity issues include the Web of Science, Google Scholar, and Transportation research in Transport Research International Documentation (TRID) (Clarivate, 2023; TRID, 2023). The information included in these sources include policy papers, design guides, toolkits, webinars, policy briefs, and case studies that address transportation equity issues (APTA, 2023; NACTO, 2023; NCMM, 2023; USDOT, 2023). Information from these sources has been used by researchers and practitioners to address transportation equity issues. Further, researchers utilize such sources for manuscript preparation for publication.

Recently, ChatGPT has been acquired by Microsoft and added to the Bing website, whereby users can chat with it and obtain text-based answers for whatever they are looking for. Like other search engines, ChatGPT can be biased, error-prone, or provide misinformation in its content. There have been some complaints about biased information from ChatGPT in various fields, such as political science, earth science, and social science (Alba, 2023; Bass, 2023; Biddle, 2023). The key difference between ChatGPT to other sources is that other search engines, such as Google search, provide the user with multiple sources of the queried information, whereas ChatGPT gives the user a single document with all the information in it. Thus, a user has no opportunity to filter misinforming content. In addition to researchers, the general community relies on search engines to obtain various information. With how easy it is to use ChatGPT, several people might be influenced to use it for day-to-day searches. Thus, the information from this source should be error-free and consistent with other trusted sources to maintain a clear understanding of various issues within the community.

Despite the debate of the biases, misinformation, and errors from ChatGPT, efforts to evaluate the extent of similarity and the content of similar and dissimilar information have still not been done to a great extent. A few studies have attempted to understand the difference between published articles and ChatGPT-generated materials. (Gao et al., 2022; Kutela, Msechu, et al., 2023). These studies concluded that ChatGPT is a good tool for academic writing but needs more human inputs to make the content more human-like. However, these studies did not provide comparable statistical analysis for the two sources of information. In this study, the abstracts of manuscripts were used to explore the difference between human-generated texts in peer-reviewed journals and ChatGPT-generated texts. The intention was to evaluate whether the information from the two sources, human and AI, was similar, and if so, to what extent and what specific content was more similar than others. Further, this study added to the body of literature regarding the methodological approach needed to explore the key difference between human-generated scientific content and ChatGPT texts. It is found that ChatGPT is still not well adept in developing correct citations. In many of the lengthy scientific writing generated by CharGPT generates either non-existing or fake citations. This study thus explores only abstracts (abstracts are citation free) to investigate the authenticity of the contents generated by ChatGPT.

The remaining sections are presented as follows. The next section presents the study methodology and discusses the data description and analytical approaches. The results and discussion section follows, then the conclusion and future studies are presented last.

METHODOLOGY

As described earlier, this study intends to explore whether ChatGPT can produce publication-ready materials comparable to human-generated text in published journals. This section presents the methodological approach used to attain the study objectives. The section is divided into two main sections, data description and analytical methods.

Data Description

To assess the similarities of the information retrieved from different sources, two types of data are necessary, human-written text and ChatGPT-generated text. In this study, authors utilized the abstract section of published papers as the human-written text.

Authors extracted transportation equity papers from the Web of Science database in which “transport equity” and “transportation equity” keywords were used to obtain the transportation equity-related manuscripts. A total of 251 manuscripts containing these keywords in their abstracts were extracted. Upon further pre-processing and duplicate checking, a total of 152 manuscripts were retained for further analysis. The abstract sections, titles, author keywords, and year of publication are the few variables of interest available in the downloaded data. Most of these papers were published in the Transportation Research Record, Transportation Research Part D-Transport and Environment, the Journal of Transport Geography, Transport Policy, and Transportation Research Part A-Policy and Practice.

To obtain the corresponding abstracts generated by ChatGPT, a prompt, which contains the directives of the actions to be taken by ChatGPT, is necessary. The following prompt was used for retrieval:

"I want you to develop write an unstructured abstract with minimum of 300 words and maximum of 500 words for publication in a scientific journal that focuses on transport equity. I will give you several titles then I want you to give me the unstructured abstract. You need to adopt a persona of high-profile researcher in transport equity, who has exceptional writing skills and has published over 100 manuscripts from various parts of the world. The abstract should have the details for at least introduction, objectives, methodology, key findings, and study implications. The first title is "Title of the paper."

The authors supplied all 152 titles to ChatGPT. The corresponding abstracts generated by ChatGPT were extracted and matched to the human-written abstracts in an Excel sheet for further analysis.

Analytical Methods

Two analytical methods, document similarity analysis and text network analysis were applied to the text data to explore the similarities and content of the transportation equity information from two sources. The document similarity analysis shows the similarity scores between the abstracts generated by ChatGPT and those written by a human. However, it does not describe the content of such abstracts. To explore the similarities and differences in the content the text network analysis was applied.

Document Similarity Analysis

Document similarity analysis is used to determine the similarities between different documents. Normally, researchers and the public consider documents to be similar if they are semantically close and describe similar concepts/themes (Selivanov, 2018; Zach, 2020).

In this study, the document is defined as the abstract. The comparison is between the abstracts written by humans to the corresponding abstracts written by ChatGPT. For each given title of the study, the human-written abstract and ChatGPT-generated abstracts are extracted and stored in the spreadsheet.

Various approaches, such as Jaccard distance, Cosine distance, Euclidean distance, and Relaxed Word Mover's Distance, can be applied to determine document similarity (Selivanov, 2018; Zach, 2020). This study used Cosine similarity with the LSA approach to determine similarities between two corresponding abstracts.

First, the documents were transformed into bag-of-words to compute the similarity between the two documents, so each document will be a sparse vector. Thus, the similarity between two documents (abstracts) can be computed as:

$$Similarity\ (doc_1, doc_2) = Cos\theta = \frac{doc_1\ doc_2}{|doc_1||doc_2|}$$

Whereby doc_1 is the human-written abstract of the paper and doc_2 is the ChatGPT-generated abstract of the same paper. The similarity score varies between 0 and 1, where a score of 1 implies that the two documents are duplicates and a score of zero implies that the two documents are not similar at all (Zach, 2020).

Text Network Analysis

Text Network Analysis (TNA) has been utilized in various fields such as literature and linguistics (Hunter, 2014), traffic safety and operations (Kutela, Das, et al., 2021; Kutela & Teng, 2021; Kwayu et al., 2021), and bibliometrics in transportation studies (Jiang et al., 2020). TNA uses nodes and edges to establish relationships between keywords within a corpus (see Figure 1), and its strength lies in its ability to visualize keywords and establish connections among them (Jiang et al., 2020; B. Kutela et al., 2021; Boniphace Kutela et al., 2021; Paranyushkin, 2011). The frequency and co-occurrence of keywords in the network are represented by the size of the nodes and the edges, respectively.

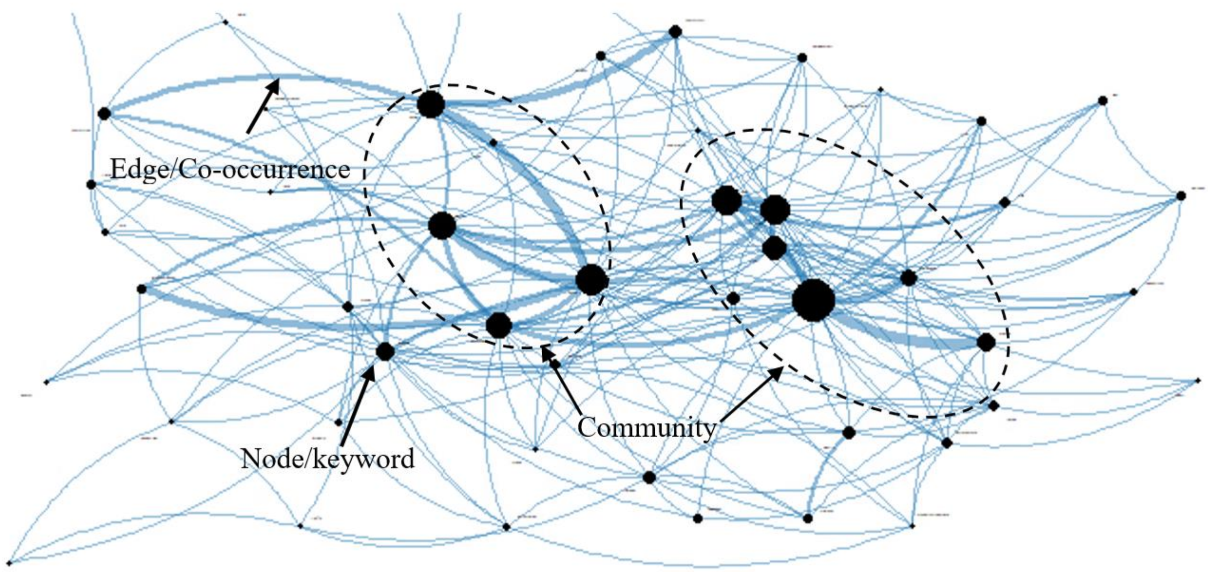


Figure 1. A Skeleton of the Text Network

Various processes are performed on the data when conducting TNA analysis. Normalization is the first process, whereby unstructured data is converted to structured data, all symbols are removed, and all texts are converted to lowercase then the output is used to create a matrix of keywords with their frequencies of occurrence (Das & Dutta, 2020; Kutela, Magehema, et al., 2022; Kutela, Novat, et al., 2022). The constructed matrix is then visualized with keywords represented as variously sized nodes based on their recorded frequencies. Various metrics can be utilized for comparative analysis, but document and collocated frequency (Kutela, Combs, et al., 2022; Kutela, Kitali, et al., 2023) are the two metrics used in this study to compare human-generated and ChatGPT-generated introductions. Document frequency is the number of documents that contain the keyword of interest. Keyword frequency, on the other hand, focuses on the number of times the keyword appears in the document. Prior to producing keyword frequency, text stemming was necessary to reduce keywords to their roots/base form. Text stemming allows for better comparison and matching of words with the same root, even if they have different suffixes (Das & Dutta, 2020). For example, "run," "runs," "running," and "runner" can all be stemmed to "run," making it easier to identify all occurrences of the root word. Collocation frequency assesses the number of times the keywords are located next to each other and offers greater insights than individual keywords do since it focuses more on the closeness between two keywords in the corpus. The collocation of keywords in a text network plays a great role in the formation of text clusters, typically referred to as a community of keywords. A community of keywords represents a group collectively clustered in the text network; there can be two or more in a text network (see Figure 1).

RESULTS AND DISCUSSION

This section presents the results and discussion. It covers the document similarity results and the text network of the abstracts.

Document Similarity Results

Table 1 presents the similarity scores between human-written and ChatGPT-generated abstracts. It is observed that the abstract is divided into six different metrics used to measure quantitative assessments. The similarity score is measured by how much they are similar on a scale of zero to one. However, some algorithms may produce negative scores, indicating dissimilarity. The similarity score increases from the minimum to the maximum category as more abstracts are developed. For example, there is a notable change between the minimum and the first quartile. On average, the content from human-written and ChatGPT-generated abstracts is 58% similar. The table also shows that the minimum similarity score is 1.4% and the maximum similarity score is nearly 97%. In the statistic, the median is 62%, in which the first quartile is 41% and the third quartile is 80%.

Table 1. Similarity Scores between Human-Written and ChatGPT-generated Abstracts

	Metrics					
	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
Similarity Score	0.01402	0.41	0.61618	0.58428	0.79738	0.96602

Although the document similarity scores present the relationship between the two documents being compared, it does not show the content of the text. Thus, text network analysis was applied to explore the pattern of the key themes of the abstracts. To perform content analysis, the documents were divided into two parts, the high and low similarity scores. The high similarity score document included all documents with similarity scores above the mean, while the low similarity score document included all documents with mean or lower similarity scores. The next section presents the content analysis results.

Text Network Results of the Abstracts

The content of the abstracts written by humans and by ChatGPT was evaluated. Figure 2 presents the text network of the human-written abstract section and ChatGPT-generated abstract section with high and low similarity scores, while Table 2 and Table 3 present the performance metrics of the text networks. The text network of human-written abstracts with high similarity scores (Figure 2 (a)) is heavily centered on the keywords *accessibility, public, transit*, *travel*, and *planning*. This is because the studies used in this paper are transportation-related. Further, the human-written abstract sections constitute keywords such as *plan, safety, environmental*, and *travel*, indicating the abstract also covers travel concerns in various studies. Some other keywords have a relatively low representation, including *time*, *regional, commuting, mobility*, and *services*, among others.

Figure 2 (b) presents the text network for ChatGPT-generated abstracts with high similarity scores. The text network is heavily centered on the keywords *accessibility*, *urban, transit*, *plan*, and *public,* similar to the human-written text network. However, the human-written text network nodes appear to be larger than those presented on the ChatGPT network. This outcome implies that these keywords appear more frequently in human-written abstract sections than in the ChatGPT-generated ones.

Table 2. Topmost Keywords from Abstracts with High Similarity Scores

Keywords						Collocations			
GPT			Human			GPT		Human	
Feature	Freq	Docfreq	Feature	Freq	Docfreq	Collocation	Count	Collocation	Count
access	431	113	access	225	67	key findings	139	public transit	39
use	288	114	use	212	105	highlights importance	55	21st century	25
find	274	151	public	211	69	public transit	66	act 21st	23
need	258	117	transit	187	51	provides valuable	59	planning process	15
includ	249	129	plan	149	64	literature review	49	flexible efficient	15
provid	245	124	travel	145	51	valuable insights	49	new york	15
polici	233	106	servic	118	46	highlights need	39	accountable flexible	15
data	224	112	polici	110	63	case studies	28	safe accountable	15
transit	216	38	urban	107	52	importance considering	51	efficient act	14
implic	215	148	model	106	44	qualitative quantitative	27	legacy users	14
plan	203	67	develop	106	66	current state	26	act legacy	14
public	202	63	analysi	106	60	conclusion provides	32	land use	14
key	196	141	research	100	58	findings indicate	66	job accessibility	13
urban	192	64	system	100	54	important implications	35	disadvantaged groups	12
import	189	117	area	94	51	implications suggest	34	travel demand	12
communiti	188	61	result	88	70	methodology includes	26	public services	12
improv	187	81	group	85	48	inform development	24	environmental review	11
research	186	89	studi	85	37	population groups	24	york city	10
servic	186	54	social	84	44	policies programs	25	autonomous vehicles	10
analysi	178	96	mobil	83	38	review existing	24	united states	10

Observing the right side of the ChatGPT-generated text network with high similarity scores (Figure 2 (b)), the linked keywords include *accessibility* and *key, low* and *income*, *communities* and *income*, and *public* and *transit.* These linked keywords can be related to transportation accessibility and the impact of income on transportation. On the other hand, the left side of the text network contains linked keywords such as *key* and *finding*, *survey* and *data*, *data* and *analysis*, *literature* and *review*, *data* and *used*, and *survey* and *methods*, which are related to analysis and recommendations/findings.

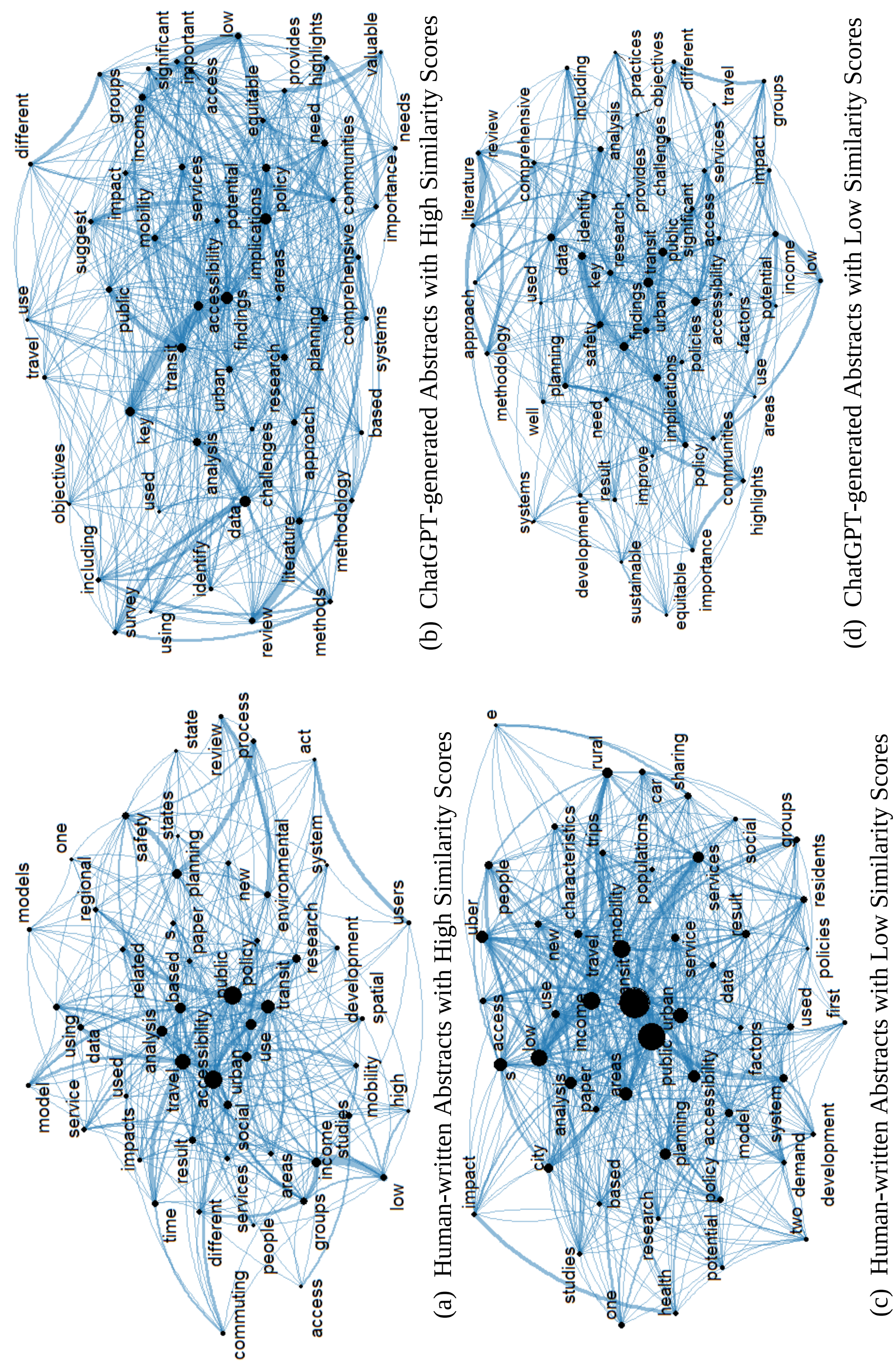
Table 3. Topmost Keywords from Abstracts with Low Similarity Scores

Keywords						Collocations			
GPT			Human			GPT		Human	
Feature	Freq	Docfreq	Feature	Freq	Docfreq	Collocation	Count	Collocation	Count
find	98	55	use	76	37	key findings	50	low income	24
implic	77	54	polici	45	30	public transit	33	public transit	22
object	67	53	public	86	28	low income	33	e scooters	9
key	74	52	result	35	28	findings indicate	27	land use	9
includ	88	48	transit	110	27	united states	25	subway accessibility	8
need	89	47	develop	42	27	provides valuable	23	public services	8
provid	92	46	citi	41	27	literature review	21	95 ci	8
data	82	43	analysi	45	26	best practices	20	irr 95	8
import	66	43	access	63	24	highlights importance	19	rural areas	7
use	100	42	servic	56	22	highlights need	19	lower income	7
highlight	57	42	incom	39	21	valuable insights	19	planning processes	7
access	163	40	improv	32	21	importance considering	18	bike sharing	7
polici	106	40	provid	24	21	ride hail	18	new york	7
signific	56	38	plan	62	20	road safety	17	transit pass	7
analysi	72	37	mobil	47	20	marginalized communities	16	last mile	7
methodolog	43	37	data	26	20	current state	15	act 21st	7
develop	59	36	system	39	19	conclusion provides	15	21st century	7
literatur	48	36	paper	23	19	mass transit	15	shared mobility	6
research	73	34	urban	41	18	mixed methods	14	york city	6
studi	40	34	area	39	18	public services	14	first last	6

Note: Docfreq = is the document frequency, which is the number of documents

Figure 2 (c) presents the text network for human-written abstracts with low similarity scores. The text network is heavily centered on the keywords *transit*, *public*, *income*, *mobility*, and *access*, similar to the text network with high similarity scores. However, the text network nodes appear to be larger than those presented on the human-written network with high similarity scores. This outcome implies that these keywords appear more frequently in human-written sections with higher similarity scores than in the sections with low similarity scores. The methodology keywords, *research*, *analysis*, and *model*, are also presented in this section.

Figure 2 (d) presents the text network for ChatGPT-generated abstracts with low similarity scores. The text network is heavily centered on the keywords *safety*, *transit*, *urban*, *planning* and *public*, similar to the human-written text network. Observing the left side of the ChatGPT-generated text network with high similarity scores (Figure 2 (d)), the linked keywords include *low* and *income*, *communities* and *income*, and *public* and *transit*. These linked keywords are similar to the text network for ChatGPT-generated abstracts with high similarity scores. On the other hand, the right side of the text network also contains linked keywords such as *key* and *finding*, *identify* and *data*, *data* and *analysis*, *literature* and *review*, *approach* and *used*, and *data* and *used*, which are related to terms of analysis and recommendations/findings.



(a) Human-written Abstracts with High Similarity Scores

(b) ChatGPT-generated Abstracts with High Similarity Scores

(c) Human-written Abstracts with Low Similarity Scores

(d) ChatGPT-generated Abstracts with Low Similarity Scores

Figure 2. Text Networks of the Content of the Abstracts

Although Figure 2 shows that the four networks share some similarities and portray some differences, a comparative analysis of the four networks can be performed using the keyword and collocation frequencies. According to the results in Table 2 in the group with high similarity scores, among the top 20 keywords, ten are common for both sides. Even for the ten common keywords, the ranking varies significantly. For instance, the keyword *transit* in ChatGPT-generated metrics is ranked ninth, appearing in 216 abstracts, while it is ranked fourth in the human-written metrics, appearing 187 times. Keywords such as *travel, model, system, group,* etc., appear only in the human-written metrics. On the other hand, the keywords *find, need, data, key,* and *community*, among others, appear only in the ChatGPT-generated metrics.

Table 3 shows similar results to the metrics generated by ChatGPT and humans with low similarity scores. There are only seven keywords common for both sides with significantly different ranks, like *polici* which is ranked third in human-written abstracts, appearing 86 times, while only ranked 13th in ChatGPT-generated abstracts, appearing 106 times. Keywords such as *find, object, key* and *include*, among others, appear only in ChatGPT-generated metrics. Keywords *result, income, transit,* and *improve*, among others, appear only in human-written metrics. This observation indicates that ChatGPT is capable of replicating some keywords.

In addition to the individual keywords, the collocated keywords results can be used to distinguish ChatGPT-generated abstracts from human-written abstracts. The results in Table 2 and Table 4 show that the two approaches differ significantly. In Table 2, which shows the metrics with high similarity scores, only one pair of collocated keywords, *public transit*, is common for both approaches. According to the results in Table 3, there are three pairs of collocated keywords that are common for both approaches with low similarity scores, *public transit, low income* and *public services.* The keywords associated with the use of the study can be observed in the collocated keywords. Such keywords include *key findings, highlight importance, conclusion provides, findings indicate*, etc. These findings indicate that the ChatGPT algorithm tends to summarize the conclusions and the possible use of the research, while that is not the case for most human-written scientific writeups. The text network and associated metrics provide the details of the content of the abstracts, which facilitates the comparison of the two.

CONCLUSIONS AND FUTURE STUDIES

This study presents a comparative analysis of transportation equity-related themes by considering the source of information. The two sources compared in this study are the traditional source, which is human-written materials, and AI-generated materials. A total of 152 highly cited papers in transportation equity were extracted from the Web of Science. A prompt was then prepared and supplied to ChatGPT to produce abstracts given the title of the manuscripts. Document similarity analysis and text network analysis, were applied to determine the relationship between human-written and ChatGPT-generated content.

It was found that on average, the human-written and ChatGPT-generated content are about 54% similar, with a minimum score of 1.4% and a maximum score of 96.7%. The content analysis of the abstracts showed a significant difference. In this case, the ChatGPT-generated materials appear to be more generic, especially when the collocation metric is considered.

Based on these findings, this study concludes that at this moment the transportation equity information retrieved through ChatGPT greatly differs from the human-written content. Therefore, researchers and the general community should exercise care when retrieving transportation equity information from ChatGPT. It is advised that researchers perform a comprehensive review of the information retrieved through ChatGPT with other human-written information. Although the two sources differ significantly, there were observed a few cases where the resulting content had high similarity scores. This is an indication that there are a few topics within transportation equity that ChatGPT performed relatively better. Thus, for such topics, the information retrieved can be used even without conducting a comprehensive review.

Although this study successfully showed the similarities and differences between the information retrieved through ChatGPT against the one written by humans, several limitations exist. First, this study utilized transportation equity-related studies. These studies are not very common for the general community, thus, the chance that their key themes were used to train ChatGPT is relatively lower. That being the case, retrieving such information using ChatGPT is relatively lower. Future studies may consider studies in areas that are common to the general community, such as social science studies. Further, this study considered the abstract section of the published papers. Abstracts tend to contain various information including the objective, methodology, and key findings. It is relatively difficult for even a human to provide an abstract when given a title. Future studies may focus on other easily generatable parts of the manuscripts given a title, such as the study objectives.

REFERENCES

Alba, D. (2023). ChatGPT, Open AI’s Chatbot, Is Spitting Out Biased, Sexist Results - Bloomberg. Bloomberg. https://www.bloomberg.com/news/newsletters/2022-12-08/chatgpt-open-ai-s-chatbot-is-spitting-out-biased-sexist-results#xj4y7vzkg.

APTA (American Public Transportation Association). (2023). Home - American Public Transportation Association. https://www.apta.com/.

Aydın, Ö., and Karaarslan, E. (2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. SSRN Electronic Journal.

Bass, D. (2023). ChatGPT maker OpenAI promises to fix the chatbot’s biases | Fortune. https://fortune.com/2023/02/16/chatgpt-openai-bias-inaccuracies-bad-behavior-microsoft/.

Biddle, S. (2023). The Internet’s New Favorite AI Proposes Torturing Iranians and Surveilling Mosques. The Concept. https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/.

Bindra, J. (2023). Will ChatGPT replace Google as our go-to web search platform? | Mint. https://www.livemint.com/opinion/columns/will-chatgpt-replace-google-asour-go-to-web-search-platform-11671733523981.html.

ChatGPT, O. A. A., and Perlman, A. (2022). The Implications of OpenAI’s Assistant for Legal Services and Society. SSRN Electronic Journal.

Clarivate. (2023). Web of Science Platform - Web of Science Group. https://clarivate.com/webofsciencegroup/solutions/webofscience-platform/.

Das, S., and Dutta, A. (2020). Characterizing public emotions and sentiments in COVID-19 environment: A case study of India.

di Ciommo, F., and Shiftan, Y. (2017). Transport equity analysis. *Transport Reviews*, 37(2), 139–151.

Frye, B. L. (2022). Should Using an AI Text Generator to Produce Academic Writing Be Plagiarism? https://papers.ssrn.com/abstract=4292283.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., and Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *BioRxiv*, 2022.12.23.521610.

Greene, T. (2023). ChatGPT will not replace Google Search | Impact of Social Sciences. https://blogs.lse.ac.uk/impactofsocialsciences/2023/01/27/chatgpt-will-not-replace-google-search/.

Hunter, S. (2014). A Novel Method of Network Text Analysis. *Open Journal of Modern Linguistics*, 4, 350–366.

Jiang, C., Bhat, C. R., and Lam, W. H. K. (2020). A bibliometric overview of Transportation Research Part B: Methodological in the past forty years (1979–2019). *Transportation Research Part B: Methodological*, 138, 268–291.

Kim, J., and Lee, J. (2023). How does ChatGPT introduce transportation problems and solutions in North America? SSRN Electronic Journal.

Kutela, B., Combs, T., John Mwekh'iga, R., and Langa, N. (2022). Insights into the long-term effects of COVID-19 responses on transportation facilities. *Transportation Research Part D: Transport and Environment*, 111, 103463.

Kutela, B., Das, S., and Dadashova, B. (2021). Mining patterns of autonomous vehicle crashes involving vulnerable road users to understand the associated factors. *Accident Analysis & Prevention*, 106473.

Kutela, B., Kitali, A. E., Kidando, E., Langa, N., Novat, N., and Mwende, S. (2023). Exploring commonalities and disparities of seattle residents' perceptions on dockless bike-sharing across gender. *City, Culture and Society*, 32, 100503.

Kutela, B., Langa, N., Mwende, S., Kidando, E., Kitali, A. E., and Bansal, P. (2021). A text mining approach to elicit public perception of bike-sharing systems. *Travel Behaviour and Society*, 24, 113–123.

Kutela, B., Magehema, R. T., Langa, N., Steven, F., and Mwekh'iga, R. J. (2022). A comparative analysis of followers' engagements on bilingual tweets using regression-text mining approach. A case of Tanzanian-based airlines. *International Journal of Information Management Data Insights*, 2(2), 100123.

Kutela, B., Msechu, K., Das, S., and Kidando, E. (2023). Chatgpt's Scientific Writings: A Case Study on Traffic Safety. SSRN Electronic Journal.

Kutela, B., Novat, N., Adanu, E. K., Kidando, E., and Langa, N. (2022). Analysis of residents' stated preferences of shared micro-mobility devices using regression-text mining approach. *Transportation Planning and Technology*, 45(2), 159–178.

Kutela, B., Novat, N., and Langa, N. (2021). Exploring geographical distribution of transportation research themes related to COVID-19 using text network approach. *Sustainable Cities and Society*, 67.

Kutela, B., and Teng, H. (2021). The Use of Dynamic Message Signs (DMSs) on the Freeways: An Empirical Analysis of DMSs Logs and Survey Data. *Journal of Transportation Technologies*, 11(01), 90–107.

Kwayu, K. M., Kwigizile, V., Lee, K., Oh, J.-S., and Oh, J.-S. (2021). Discovering latent themes in traffic fatal crash narratives using text mining analytics and network topology. *Accident Analysis and Prevention*, 150, 105899.

Martens, K., Bastiaanssen, J., and Lucas, K. (2019). Measuring transport equity: Key components, framings and metrics. *Measuring Transport Equity*, 13–36.

Mobility-Innovators. (2023). How ChatGPT will be game changer for Public Transit agencies? - Mobility Innovators. https://mobility-innovators.com/how-chatgpt-will-be-game-changer-for-public-transit-agencies/.
NACTO. (2023). National Association of City Transportation Officials | National Association of City Transportation Officials. https://nacto.org/.
NCMM. (2023). Home - National Center for Mobility Management. https://nationalcenterformobilitymanagement.org/.
Noever, D., and Ciolino, M. (2022). The Turing Deception. ArXiv.
Paranyushkin, D. (2011). Identifying the Pathways for Meaning Circulation using Text Network Analysis. *Venture Fiction Practices*, 2(4). www.noduslabs.com.
Qadir, J. (2022). *Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education*.
Selivanov, D. (2018). Documents similarity. https://text2vec.org/similarity.html.
TRID. (2023). About TRID | Information Services. https://www.trb.org/InformationServices/AboutTRID.aspx.
USDOT. (2023). Equity | US Department of Transportation. https://www.transportation.gov/priorities/equity.
Yalalov, D. (2022). 100 Best ChatGPT Prompts to Unleash AI's Potential | Metaverse Post. https://mpost.io/100-best-chatgpt-prompts-to-unleash-ais-potential/.
Zach. (2020). How to Calculate Jaccard Similarity in R. https://www.statology.org/jaccard-similarity-in-r/.