ORIGINAL ARTICLE

LEARNED PUBLISHING

(wileyonlinelibrary.com) doi: 10.1002/leap.1647 Received: 2 April 2024 | Accepted: 15 November 2024 | Published online in Wiley Online Library: 1 December 2024

'AI-navigating' or 'AI-sinking'? An analysis of verbs in research articles titles suspicious of containing AI-generated/assisted content

Ruben Comas-Forgas [1*], Alexandros Koulouris [2] and Dimitris Kouis [2]

[1]Department of Applied Pedagogy and Educational Psychology, University of the Balearic Islands, Palma de Mallorca, Spain

[2]Department of Archival, Library & Information Studies, University of West Attica, Egaleo, Greece

ORCID:
R. Comas-Forgas: 0000-0002-8885-753X
A. Koulouris: 0000-0002-4011-2678
D. Kouis: 0000-0002-5948-9766

*Corresponding author: Ruben Comas-Forgas, University of the Balearic Islands, Palma de Mallorca, Spain.
E-mail: rubencomas@uib.es

Abstract: Our study investigates the impact of Artificial Intelligence (AI), specifically generative AI technologies (GAI), on the linguistics of academic article titles. Triggered by suspicious of increased usage of specific verbs in article titles, this research hypothesizes that GAI tools may be influencing the language of scientific communication. To explore this hypothesis, we conducted a comprehensive analysis on the frequency and distribution of 15 selected verbs in research article titles, using data extracted from the SCOPUS database spanning 2015 to 2024. The methodology integrates qualitative observations with a bibliometric approach, examining the presence and trends of these verbs across multiple scientific disciplines. The findings reveal a marked increase in these verbs, pointing towards AI's involvement in title generation. We also explore document characteristics, such as disciplinary backgrounds and publication contexts, to gauge AI's impact on academic writing. Furthermore, the research attempts to quantify the extent of AI-assisted title generation. Despite several limitations, this investigation paves the way for future studies to broaden the linguistic and database scope. It underscores the need for establishing AI usage standards in academic publishing, contributing valuable insights into the ongoing dialogue about AI's integration into academic writing.

Keywords: artificial intelligence, databases, scientific communication, titles

INTRODUCTION

The origin of this research paper traces back to an intriguing observation made by the corresponding author, who, in his role as an active reviewer for journals specializing in education and information sciences, noted a peculiar trend during the year 2023. The term 'navigating' emerged with surprising frequency in the titles of articles submitted for review, a pattern that was both unusual and thought-provoking. This linguistic anomaly, particularly the recurrent use of 'navigating', sparked a suspicion that led to a deeper inquiry: could these titles be indicative of a broader trend towards AI-generated academic content?

The hypothesis that AI might play a significant role in shaping the discourse within scientific papers fuelled the design and

development of this research. The suspicion was not unfounded, considering the advancements in AI and its increasing incorporation into academic writing tools (Golan et al., 2023; Lund et al., 2023). AI's capabilities in producing coherent and contextually relevant content have raised both opportunities and challenges within the scholarly community (Dergaa et al., 2023). The peculiar repetition of the verb 'navigating' in article titles reviewed by the corresponding author served as a catalyst for exploring the extent to which GAI-generated content has permeated academic literature.

The competence of academic writing, a cornerstone skill within scientific contexts, has undergone notable changes attributed and encompassed to technological innovations (Han et al., 2021). The advent of diverse digital instruments designed to support scholars through their investigative, drafting, and composition phases evidenced a significant shift in academic writing practices (Curry, 2023; Godoy, 2020). These innovations, involving everything from digital repositories and collaborative online environments to tailor-made authoring applications, have revolutionized the characteristics, modes of production and distribution of academic manuscripts and scientific communication in general (Strobl et al., 2019).

This evolution in academic writing practices has been further accelerated since the last trimester of 2022, perhaps even more disruptively and with deeper implications, by the rapid and widespread penetration of Large Language Models (LLMs) and GAI technologies (Májovský et al., 2023). Subsequently, the utilization of these means in scholarly communications has been the subject of considerable debate (Bin-Nashwan et al., 2023). Initial optimism regarding their potential to substitute for most types of formal writing, including proposals to recognize them as authors in academic publications (Dwivedi et al., 2023; Stokel-Walker, 2023), has given way to a more refined comprehension of their strengths and significant weaknesses (Lingard, 2023; Meyer et al., 2023).

There is no doubt that integration of AI tools into scientific research is on the rise, signalling a transformative shift in research methodologies, as indicated by a Nature survey involving over 1600 global researchers (Van Noorden & Perkel, 2023). The study highlights a growing consensus among scientists that AI will play a critical or essential role in their fields within the next decade, reflecting the anticipated centrality of these tools in future research practices. Furthermore, this massive survey identifies that investigators have explored the capabilities of ChatGPT, with approximately 30% of scientists having employed generative AI technologies for the purpose of composing research papers.

Evidencing the impact of GAI at a broad scale presents considerable challenges; nowadays it is almost impossible giving reasonable figures on how much of the existing scientific literature it is produced or supported by systems as Chatgpt, Bard or Copilot. Instances where papers produced entirely by Large Language Models (LLMs) have been detected, often due to the inclusion of distinctly non-human phrases like 'As an AI language model…', have been reported, though they remain relatively infrequent when contrasted with the volume of several million papers published annually. It has been recognized, however, that the omission of such identifiable phrases may complicate the task of straightforward detection (Gray, 2024).

Key points

- Observed surge in 'navigating' within titles hinted at GAI's influence on academic lexicon.
- AI technologies like ChatGPT are reshaping scientific discourse, blending creativity with potential standardization.
- The study reveals a significant year-over-year growth in specific verbs across academic titles, suggesting GAI assistance.
- Findings urge a balance between GAI's benefits in enhancing writing and risks of diminishing scholarly originality.

Liang et al. (2024) have uncovered compelling evidence indicating the utilization of ChatGPT and analogous technologies by researchers to produce peer reviews for conference papers within the artificial intelligence domain. Remarkably, their analysis did not reveal a comparable trend in the peer reviews for journals under the Nature portfolio. Despite the intricate and advanced nature of their research, Liang *et al*. managed to identify surprisingly straightforward patterns. Specifically, they observed that certain adjectives were employed with significantly greater frequency—ten to thirty times more often—in reviews penned in 2023, subsequent to the public launch of ChatGPT 3.5, compared to usage rates in prior years.

In 2023, an analysis (Gray, 2024) revealed a notable and uneven surge in the frequency of certain keywords, both individually and collectively. It is speculated that a minimum of 60,000 publications (accounting for just above 1% of total articles) received assistance from LLMs. This figure could potentially be adjusted and further detailed through the examination of additional paper attributes or the discovery of more keywords suggestive of LLM involvement.

In this investigation, following the patterns of Gray (2024), our analytical attention is directed on research manuscript titles, an approach underpinned by both pragmatic and theoretical considerations. Given that LLM-enabled platforms like ChatGPT (GPT4 Plus) are limited in the volume of text they can generate and are more susceptible to generating erroneous or fabricated content in lengthier compositions (Zheng & Zhan, 2023), we considered focusing our object of analysis towards examining titles. The existing corpus of evidences examining the structure and rhetoric of research manuscript titles and abstracts suggests a high degree of rhetorical and stylistic uniformity (Fox & Burns, 2015; Samar et al., 2014). These characteristics make titles an interesting and significant subject for assessing the potential influence of AI-driven text generation or assistance. Titles hold significant value within the scholarly communication ecosystem, serving as a pivotal element in the editorial screening process, influencing reviewers' willingness to engage with a manuscript,

and playing a crucial role in the dissemination and visibility of research findings (Sagi & Yechiam, 2008). Often, the title, along with the abstract, is one of the few pieces of content readily accessible outside of subscription-based barriers, emphasizing its importance in the broader academic impact. All these elements are considered to justify focusing our approach on the titles of scientific manuscripts as are the first and more informative element of scientific communication.

The research objectives for a study on the analysis of action verbs in research articles, with a focus on identifying potential indicators of GAI-generated or assisted content, particularly text produced by ChatGPT, are outlined as follows:

- Identify suspicious verbs: to systematically determine the set of action verbs that are recurrently utilized by ChatGPT in proposing titles for scientific manuscripts. This involves identifying verbs that may signal the involvement of generative GAI technologies in the suggestion of article titles, thereby highlighting patterns that differentiate GAI-assisted content from human-generated texts.
- Trend analysis of verb usage: to conduct a longitudinal analysis of the trend concerning the usage of these identified action verbs in scientific communication over the past decade, with a projection for 2024. This involves quantitatively tracking the frequency and distribution of these verbs in academic titles across various disciplines, thereby elucidating any shifts or notable trends in linguistic practices within scholarly communication that may correlate with the advent and evolution of GAI writing aids.
- Characteristics of the documents with suspicious GAI-generated or assisted titles: to analyse the characteristics of research outputs that have incorporated the identified suspicious verbs in their titles over the last 15 months. This analysis will help understanding the broader context and potential implications of GAI-assisted academic writing, including the scope of GAI's influence on various fields and the academic community's reception and integration of GAI-generated content.
- Quantification of AI-assisted manuscripts: to estimate the potential number of scientific manuscripts that have utilized ChatGPT for title generation. This objective seeks to quantitatively assess the extent to which GAI, specifically ChatGPT, has been employed in the creation of titles for scientific articles. By developing a methodology to approximate the volume of titles influenced by ChatGPT, the study aims to provide a clearer picture of the penetration and impact of GAI technologies on the titling process in academic writing.

Through these research objectives, the study aims to set-up evidence on the ways in which generative AI technologies, particularly ChatGPT, are influencing academic writing practices. By examining the linguistic patterns and the broader characteristics of potentially GAI-assisted research articles, the study seeks to contribute valuable insights into the fast-changing panorama of scientific communication in the age of AI.

MATERIAL AND METHODS

Step 1: Identify potential suspicious GAI-generated/assisted frequently used verbs

An initial list of verbs considered being suspiciously frequent in GAI-generated or assisted titles was compiled based on the co-author's observations during the review of numerous article submissions during the last years. These verbs include: navigating, enhancing and leveraging.

To expand upon this preliminary list, an empirical approach was employed involving 40 abstracts sourced from recent articles published across four distinct fields in SCOPUS: social sciences, sciences and engineering, humanities and health sciences, with 10 abstracts aleatory selected from each discipline. These abstracts served as inputs to ChatGPT4.0, with the prompt: 'Please, based on this abstract, redact a title for an academic/research paper'. For each abstract, 5 regenerations were executed in ChatGPT4.0. which makes a total of 200 titles generated by the GAI model (see Supplementary material 1).

Special attention was paid to verbs, in present particle form, that appeared as the initial word in the AI-generated titles, leading to the identification of additional verbs deemed suspicious. This process culminated in a curated list of 25 key verbs anticipated to be indicative of AI-generated content.

Upon identifying a suspicious verb in the participle present form, the researchers implemented a rigorous initial screening process. This process entailed conducting searches for the presence of each identified verb within the titles of articles catalogued in the SCOPUS database. The primary objective of this screening was to ascertain any significant variations in the usage volume of these verbs from the year 2023 onward. A significant increase in the frequency of a particular verb during this period, as compared to historical data, was interpreted as an indicator of potential GAI generation or assistance. This criterion for determining suspicion was predicated on the assumption that a marked rise in the use of specific action verbs could be attributed to the growing influence and application of GAI technologies, such as ChatGPT, in the academic writing process.

This enhanced screening step was decisive in refining the list of verbs earmarked for further analysis. Verbs that exhibited a notable uptick in usage within the specified timeframe were earmarked as 'suspicious of being GAI-generated'. This methodological augmentation ensured a data-driven approach to identifying verbs most likely influenced by GAI interventions in scientific manuscript title generation.

By integrating this empirical validation step into the methodology, the study leverages quantitative data from a reputable academic database to corroborate the initial qualitative observations. This dual-faceted approach strengthens the validity of the identified list of suspicious verbs, setting a robust foundation for subsequent analyses of GAI's linguistic footprint in academic title generation.

Step 2: Conducting a bibliometric analysis

The study followed a detailed analysis of the presence and evolutionary trend of the previously identified suspicious verbs within the corpus of literature indexed in SCOPUS, Web of Knowledge and DOAJ. Researchers conducted documentation search individually per each of the 15 suspicious AI-generated or assisted verbs and joint searches by generating a single search equation with all verbs. The investigation has been strategically timed to include data up to March 31st, allowing for the inclusion of early 2024. By extrapolating the frequency of these verbs within the first trimester of 2024 and projecting this across the entire year (multiplication by four), the study aimed to capture the most current linguistic trends and potentially amplify the visibility of any significant findings during 2024. A trend analysis was conducted to determine the exact increase in the usage of the identified verbs in titles. Using data from SCOPUS, researchers conducted an analysis of the characteristics of the manuscripts.

Data process and analysis was done by generating a data matrix in Excel file and calculating the diverse operations by using the Data Analyst, a chatbot built by OpenAI that can be accessed via ChatGPT4.

ANALYSIS AND RESULTS

Verbs suspicious of being part of GAI-generated manuscript titles

The investigation into the linguistic characteristics indicative of ChatGPT involvement in the generation of scientific manuscript titles culminated in the identification of a distinct set of 15 verbs. These verbs were observed to exhibit a notable increase in their utilization within the span of 2023–2024, suggesting their potential origin from AI-generated or assisted content, particularly in the context of manuscript titles. The comprehensive analysis aimed to pinpoint verbs whose frequency surged in a manner that deviated from previous trends, thereby hinting at the influence of AI technologies like ChatGPT in shaping academic discourse. The list of verbs identified and the number of articles (NA), indexed in SCOPUS, that used them in the last 10 years and its year-over-year growth rate (YoYG) can be consulted in Table 1.

Based on the data compiled in Table 1 and represented in Fig. 1 (only illustrating the last 5 years and limiting the YoYG representation up to 100% in order to make the image more clear), which showcases the YoYG rates of documents indexed in SCOPUS featuring specific verbs in their titles from 2020 to 2024, we can draw several conclusions. The data from 2023 and 2024 is particularly noteworthy, as the 'suspicious' verbs identified in our study exhibit steep increases in YoYG, suggesting an elevated usage in manuscript titles that could be associated with GAI assistance or generation. The mean average YoYG rate for indexed documents in SCOPUS with one of the verbs taken into consideration in our analysis for the years 2023 and 2024 raises up to 99.9%.

The mean YoYG rates for 2023–2024 highlight a range of verbs with varying degrees of increase, with some showing spectacular increasing (see Fig. 2). We generated the following classification based on the data:

- High YoYG (≥100%): Revolutionizing (381%), Unleashing (208%), Unlocking (137.20%) and Unveiling (157%) have the highest mean YoYG rates, suggesting a strong correlation with AI's influence in constructing titles that suggest a sense of novelty and breakthrough.
- Moderate YoYG (≥50% ≤ 100%): Verbs like Unravelling (88.75%), Advancing (65.55%), Harnessing (63.15%), Enhancing (62.90%), Navigating (73.55%), Uncovering (70.65%) and Pioneering (54.10%) show significant but more moderate increases when compared to the previous.
- Low YoYG (<50%): Deciphering (43.95%), Exploring (41.55%), Leveraging (28.50%) and Bridging (23.75%) demonstrate relatively lower YoYG rates, which may indicate a more stable use of these terms over time and a less pronounced influence of AI text generation tools.

Characteristics of the documents

When, on March the 31st, applying the following search equation into SCOPUS database 'TITLE (unravelling OR harnessing OR unraveling OR unlocking OR uncovering OR advancing OR unleashing OR revolutionizing OR exploring OR pioneering OR deciphering OR bridging OR navigating OR enhancing OR unveiling OR leveraging) AND PUBYEAR > 2023 AND PUBYEAR < 2024' a total of 84,366 (2023 = 69,459 and 2024 = 23,907) documents have been listed, that after calculating an estimation for the whole 2024 (by multiplying X 4 the number of documents returned as for the first trimester) represents a total of 165,087 documents indexed in SCOPUS with titles in which, at least, one of the 15 analysed verbs appear in the title.

Figure 3 illustrates the YoYG by type of document indexed in SCOPUS comparing the 2021–2022 period with the 2023–2024 estimation. The remarkable growth observed in categories such as 'Letter' and 'Review' points towards a potential growing acceptance and reliance on AI for drafting concise communications and comprehensive literature reviews. This trend might reflect the utility of AI in assisting researchers to efficiently manage the vast amounts of literature available and to succinctly communicate complex ideas or findings. The high growth rate in 'Letters' could indicate an increased use of AI tools for rapid, focused communication, possibly in response to the dynamic nature of certain fields where timely dissemination of findings is crucial. Similarly, the growth in 'Editorial' and 'Short Survey' documents may demonstrate the effectiveness of AI in aiding the synthesis of research trends and perspectives and redacting attractive and 'click bait' style titles.

TABLE 1 List of suspicious verbs used in AI-generated or assisted titles (NA and YoYG).

Verb	Data	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Navigating	NA	470	516	596	699	713	938	1106	1240	2043	3724
	YoYG	X	9.8%	15.5%	17.3%	2%	31.6%	18%	12.1%	64.8%	82.3%
Enhancing	NA	4453	4982	5428	6128	7060	7800	8519	9518	15,111	25,228
	YoYG	X	11.9%	9%	12.9%	15.2%	10.5%	9.2%	11.7%	58.8%	67%
Unveiling	NA	260	256	332	366	473	668	794	1060	2650	7000
	YoYG	X	−1.5%	29.7%	10.2%	29.2%	41.2%	18.8%	33.5%	150%	164.1%
Leveraging	NA	791	892	1004	1150	1386	1631	2071	2347	3471	3788
	YoYG	X	12.8%	12.6%	14.5%	20.5%	17.7%	27%	13.3%	47.9%	9.1%
Unravelling	NA	679	759	812	954	1136	1359	1613	1885	2944	6516
	YoYG	X	11.8%	7%	17.5%	19%	19.6%	18.7%	16.9%	56.2%	121.3%
Unlocking	NA	178	217	242	254	292	309	411	443	1186	2452
	YoYG	X	21.9%	11.5%	5%	15%	5.8%	33%	7.8%	167.7%	106.7%
Uncovering	NA	371	421	438	503	542	660	791	880	1390	1869
	YoYG	X	13.5%	4%	14.9%	7.7%	21.8%	19.9%	11.3%	106.8%	34.5%
Advancing	NA	815	820	1011	1054	1144	1174	1380	1529	2587	4188
	YoYG	X	0.6%	23.3%	4.6%	8.5%	2.6%	17.6%	10.8%	69.2%	61.9%
Unleashing	NA	41	50	60	62	69	72	81	110	440	948
	YoYG	X	21.9%	20%	3.3%	11.3%	4.3%	12.5%	35.8%	300%	115.4%
Revolutionizing	NA	37	41	41	37	55	56	71	80	616	1180
	YoYG	X	10.8%	0%	−9.7%	48.6%	1.8%	26.8%	12.7%	670%	91.56%
Harnessing	NA	363	416	529	553	589	771	877	915	1525	2436
	YoYG	X	14.6%	27.2%	4.5%	6.5%	30.9%	13.7%	4.3%	66.6%	59.7%
Exploring	NA	6923	7725	8562	9537	10,814	12,455	14,419	16,507	23,632	33,096
	YoYG	X	11.6%	10.8%	11.4%	13.4%	15.2%	15.8%	14.5%	43.1%	40%
Pioneering	NA	135	167	142	148	148	136	212	211	265	484
	YoYG	X	23.7%	−14.8%	4.2%	0%	−8.1%	55.8%	−0.5%	25.6%	82.6%
Deciphering	NA	330	374	388	462	598	660	863	1041	1240	2092
	YoYG	X	13.3%	3.7%	19%	29.4%	10.4%	30.8%	20.6%	19.2%	68.7%
Bridging	NA	1410	1395	1421	1440	1618	1636	1696	1912	2335	2928
	YoYG	X	−1.1%	1.9%	1.3%	12.4%	1.1%	3.7%	12.7%	22.1%	25.4%

The YoYG rates reflect significant growth across all disciplines from 2021–2022 to 2023–2024 (see Fig. 4), with the 'Multidisciplinary' category experiencing the highest growth rate at 185%, indicating a substantial increase in manuscripts with suspicious AI-generated or assisted titles. The 'Sciences and Engineering' and 'Health Sciences' categories also saw remarkable growth, highlighting the three critical areas with increasing YoYG rates over 100%. The growth rates in 'Social Sciences' and 'Arts and Humanities' were more moderate but still substantial, evidencing a steady increase in potential AI-generated or assisted titles in these fields.

The growth rates (when comparing data for 2021–2022 with estimated data for 2023–2024) per countries of the authors −59.8% for authors from English speaking countries and 144.5% for authors from non-English speaking countries- highlight a marked increase in the potential use of AI tools in academic research and writing processes globally, with a notably sharper ascent in the latter group (see Fig. 5). This data suggests a rapid and solid adoption and integration of AI technologies in academic writing assistance across non-English speaking regions, perhaps reflecting a growing interest and need in applying GAI to overcome language barriers. The relatively lower, yet substantial, growth

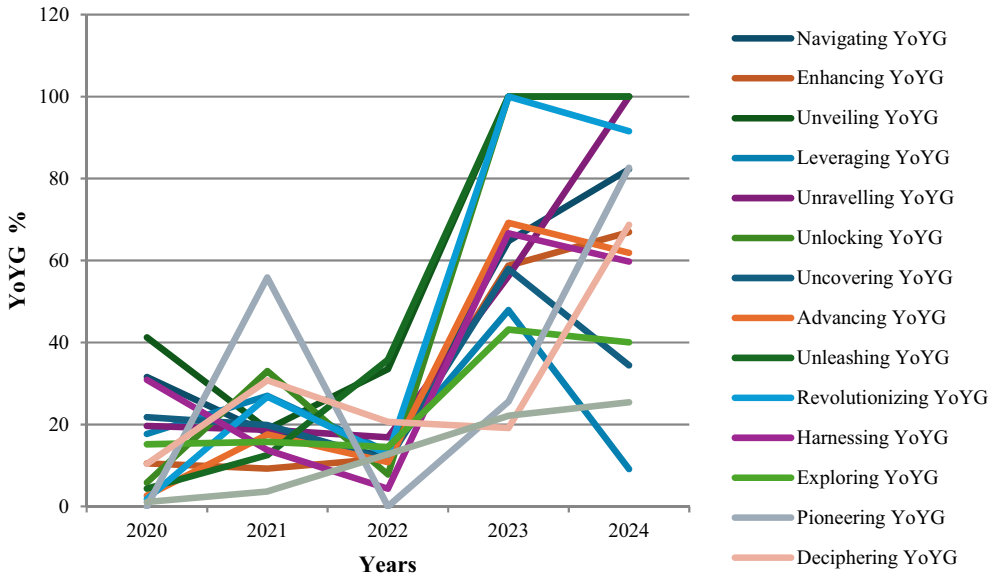

FIGURE 1 YoYG rates of the 15 verbs analysed (2020–2024).

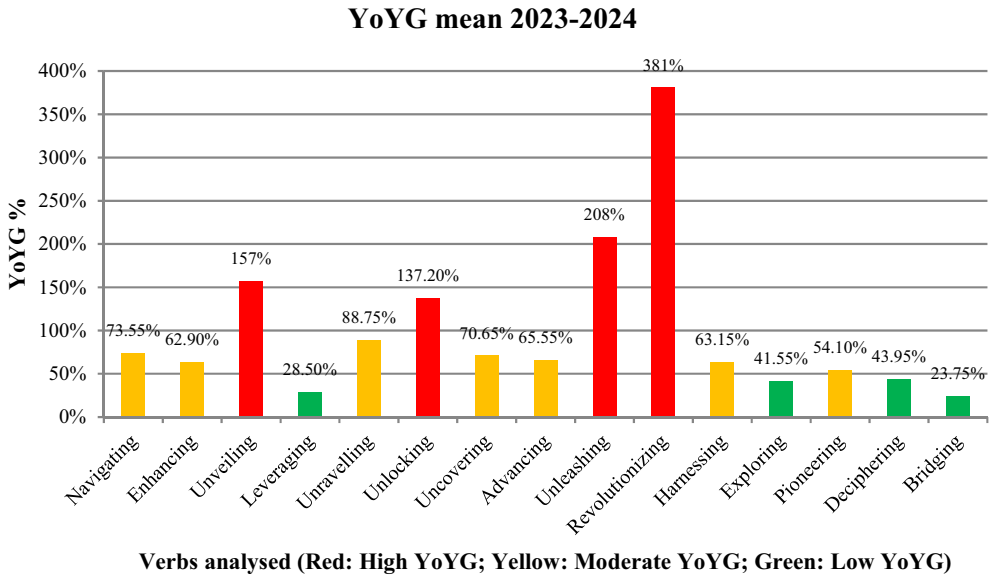

FIGURE 2 YoYG mean for each verb (2023–2024).

rate among English-speaking countries indicates a potentially steady, ongoing integration of AI in scientific communication in these regions as well, although the starting base may have been higher or the adoption more gradual.

The prevalence of potentially GAI-generated titles in academic manuscripts

To calculate the potential number of titles generated or assisted by GAI, we followed the following steps: (a) calculate the mean YoYG for 2016 to 2022; (b) estimate the YoYG for 2023 and 2024 based on the mean of YoYG 2016–2022. Assuming the trend without external influence (such as AI assistance) would continue similarly, we estimate the YoYG for 2023 and 2024 using the mean YoYG calculated in step a. This estimation assumes the presence of certain features in manuscript titles (in our case the analysed verbs) would follow the historical growth rate; (c) calculate the potential AI-generated or assisted titles; to estimate the potential number of AI-generated or assisted titles, we first estimate the expected number of manuscripts for 2023–2024 based on the historical YoYG rate and the difference between this estimation and the real figures result

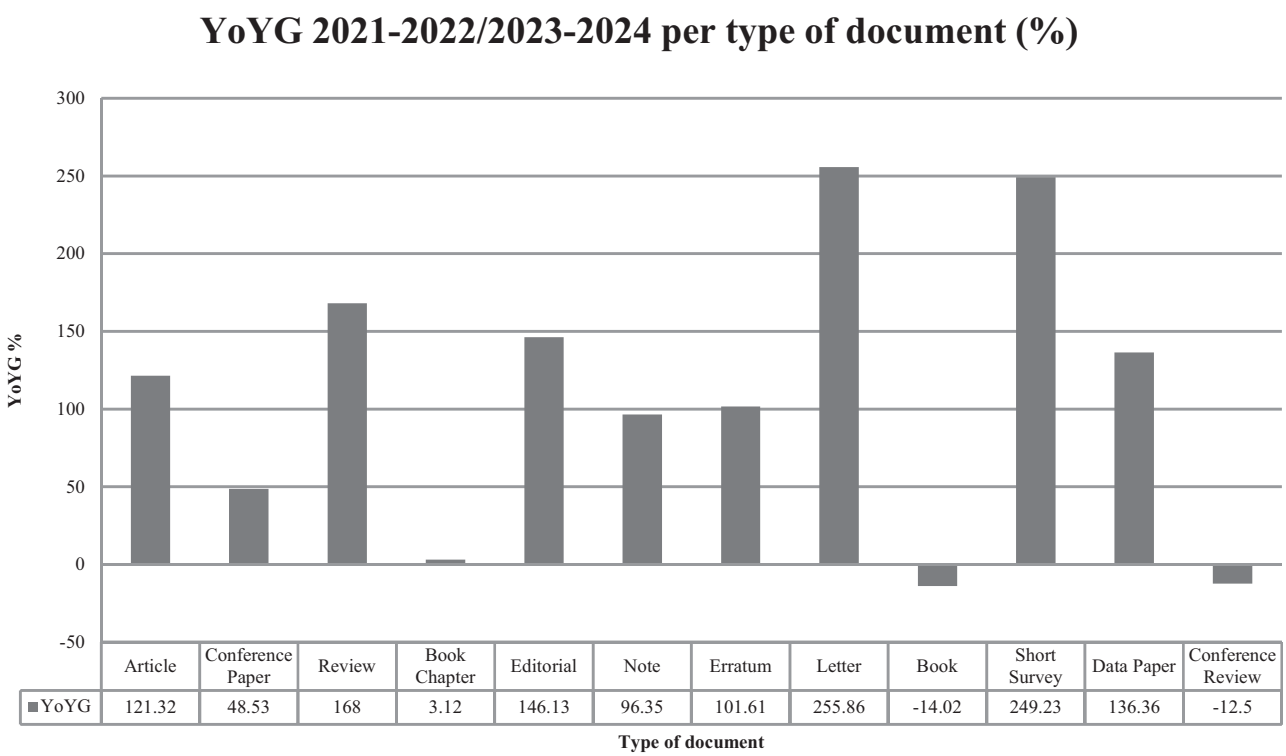

FIGURE 3 YoYG by type of document (21–22 *vs.* 23–24).

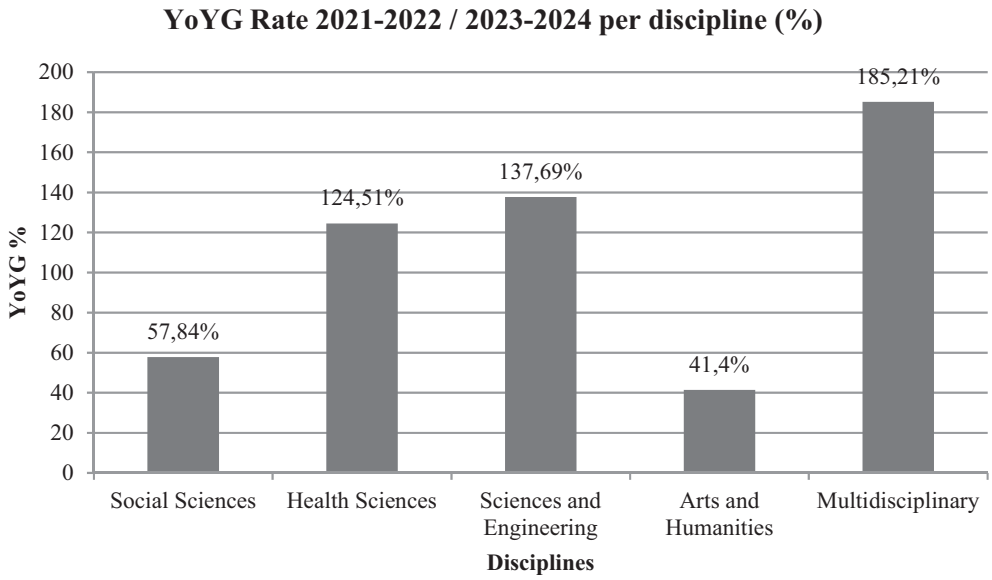

FIGURE 4 YoYG by discipline (21–22 *vs.* 23–24).

the number of potentially GAI-generated or assisted manuscript titles.

The dataset (see Fig. 6 and Table 2) indicates a notable surge in manuscript titles within SCOPUS that are suspicious of being AI-generated or assisted, suggesting a significant turn in academic publishing. With a combined total of 63,780 such titles over 2 years (considering the 15 analysed verbs in the titles), there is a discernible impact of GAI on the scholarly domain. The substantial year-over-year increase in titles featuring verbs like 'Enhancing' and 'Exploring' could reflect a GAI-driven influence on the lexicon of academic titles, pointing to a possible standardization or trend in the terminology used in scholarly articles. This trend

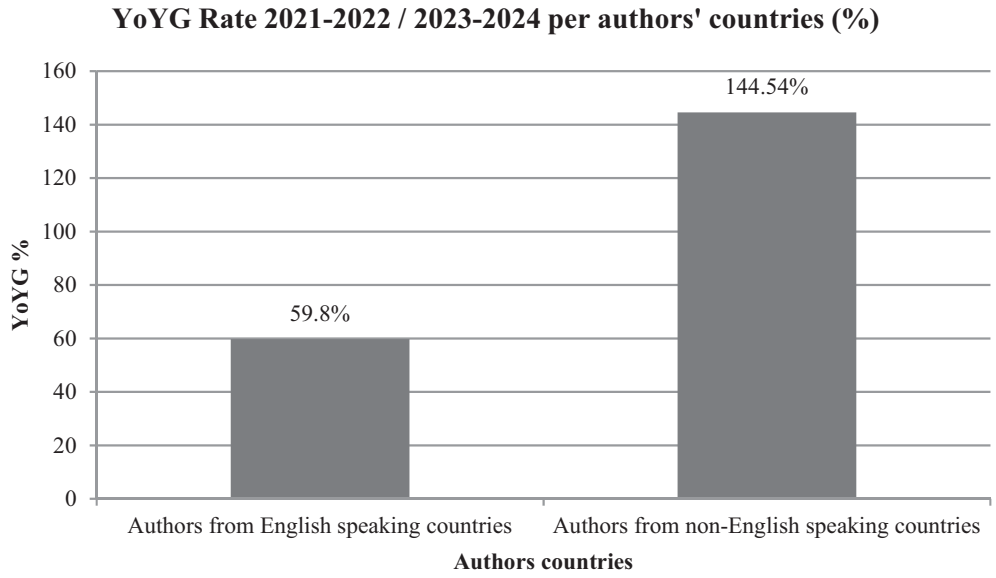

FIGURE 5 YoYG by countries (21–22 *vs.* 23–24).

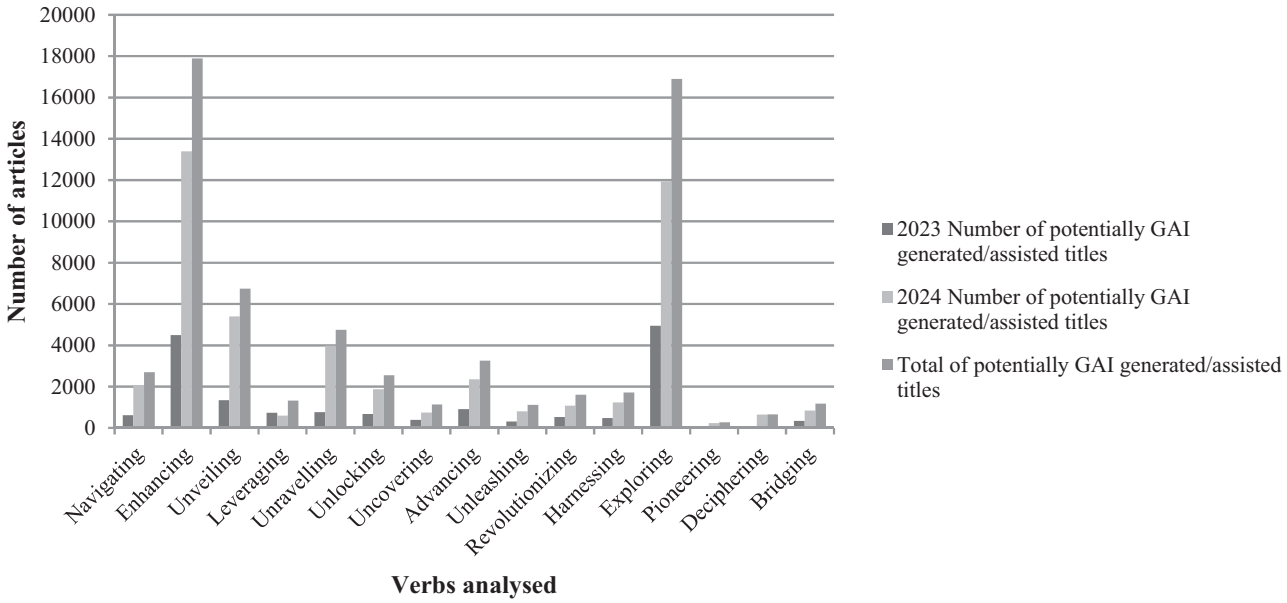

FIGURE 6 Prevalence estimation of potentially GAI-generated titles (2023–2024).

raises important questions about the originality and authenticity of academic work in the era of advanced GAI technologies.

DISCUSSION

The striking uptrend in certain verbs during the last 2 years evidenced in our study potentially reflects a broader integration of GAI in the research publication process. The recurrent use of these specific verbs may not just be a stylistic choice but the evidence of AI assistance in academic writing. This trend raises important questions about the originality of research titles and the need for careful evaluation of GAI's expanding role in academic writing. As the academic community grapples with these developments, the establishment of standards and guidelines for AI's use in research communication becomes imperative to maintain the authenticity of scholarly output. The remarkable growth in the use of specific verbs identified as potentially AI-generated or GAI-assisted in manuscript titles from 2023 to 2024 suggests a profound AI influence on academic lexicon, a trend that has not been extensively documented in previous literature. From the scarce existing corpus of evidence, recent studies have similarly

TABLE 2 Prevalence estimation of potentially GAI-generated titles (2023–2024) and YoYG rate.

Verbs	2023 Number of potentially GAI generated/assisted titles	2024 Number of potentially GAI generated/assisted titles	Total of potentially GAI generated/assisted titles	YoYG (2023–2024)
Navigating	615	2080	2695	238.2%
Enhancing	4499	13,395	17,894	197.6%
Unveiling	1347	5397	6744	300.6%
Leveraging	733	592	1325	−19.2%
Unravelling	762	3989	4751	423.6%
Unlocking	680	1874	2554	175.5%
Uncovering	393	740	1133	88.3%
Advancing	910	2348	3258	157.3%
Unleashing	313	801	1114	155.9%
Revolutionizing	526	1078	1604	104.9%
Harnessing	478	1237	1715	158.5%
Exploring	4947	11,944	16,891	141.4%
Pioneering	36	236	272	555.5%
Deciphering	11	641	652	5727.2%
Bridging	337	841	1178	149.5%
Total	16,587	47,193	63,780	184.4%

highlighted the encroaching influence of generative AI technologies on various facets of academic writing. The findings of our investigation resonate with those of Liang et al. (2024), who demonstrated the use of ChatGPT and similar technologies in producing peer reviews. Additionally, a study by Gray (2024) on the prevalence of AI-generated content in scientific publications underlines the growing penetration of AI tools like ChatGPT in scholarly communication. Another recent study conducted by Haider et al. (2024) in which authors retrieved and analysed a sample of scientific papers with signs of GPT-use; the results reveal that, approximately, two-thirds of the retrieved papers were discovered to have been generated, at least partially, through the undisclosed and potentially misleading use of GAI. Following the same pattern, Kobak et al. (2024) estimated that at least 10% of biomedical abstracts in the first half of 2024 had been written using large language models (LLMs), translating to around 150,000 papers annually. The team analysed 14 million abstracts from PubMed, covering publications between 2010 and June 2024. Their analysis found a link between LLM usage and an increase in stylistic terms like 'delves', 'showcasing', and 'underscores'. These patterns helped them estimate the extent of GAI usage in scientific writing. Finally, another study (Uribe & Maldupa, 2024) assessed the influence of ChatGPT on dental research writing. Analysing 299.695 abstracts from PubMed (2018–2024), researchers found a significant increase in the use of 'signalling words' indicative of ChatGPT. The frequency of such words rose from 47.1 to 224.2 per 10,000 papers, with 'delve' showing the highest increase. Our findings extend this discourse, suggesting that GAI technologies may be contributing to evolving linguistic norms in scientific communication, possibly influencing the rhetoric and stylistic uniformity observed in recent years.

This study demonstrates GAI's linguistic influence in academic writing by analysing 15 verbs in scientific manuscript titles, highlighting both benefits and limitations. The narrow focus was necessary to manage scope but limits generalizability to the broader AI-generated academic lexicon. Future research will expand the analysis to more linguistic markers for a deeper understanding of GAI's patterns. Additionally, the exclusive use of SCOPUS data limits data diversity. Future studies will incorporate additional databases for a more comprehensive view of GAI's impact on academic writing.

A second limitation is that our study estimates 2024 data trends based on first-trimester observations. While this offers valuable insights, it carries inherent uncertainty. Predictive modelling over a short timeframe may not fully account for the dynamic nature of academic publishing. Though conducted with a robust framework, we advise interpreting these forecasts cautiously as one of several possible outcomes. Additionally, we did not examine individual articles for explicit author disclosures on GAI tools like ChatGPT, limiting insights into transparency. Future research will explore how such disclosures affect academic integrity. These limitations guide further research into AI's changing role in scholarly communication, which we aim to address in future studies. The findings of this study suggest a notable and growing influence of Generative AI (GAI) on academic writing, particularly in

the titles of research papers. Through the analysis of 15 selected verbs, we observed significant increases in the usage of certain verbs from 2023 to 2024, indicating a potential correlation between these linguistic trends and GAI-assisted title generation. This trend is evident across various disciplines and document types, underscoring the widespread nature of GAI's impact on academic communication.

Our study demonstrates that the presence of specific verbs such as 'Revolutionizing', 'Unveiling', and 'Enhancing' in academic titles has surged dramatically in recent years. These verbs, often indicative of breakthrough or novelty, may be favoured by GAI tools like ChatGPT when generating or assisting in the creation of titles. This suggests that GAI is not only being used to assist in writing but is also shaping the stylistic choices in academic discourse, reflecting both creativity and, in some cases, a trend towards uniformity.

The quantitative results, especially the Year-over-Year Growth (YoYG) rates, support the conclusion that GAI is increasingly influencing academic title and text generation. The data reveals that certain verbs exhibited growth rates exceeding 500%, which starkly contrasts with the average linguistic shifts typically seen in academic writing. This steep rise in verb usage suggests that GAI has become a significant force in shaping the language of research publications. Moreover, the trend is particularly evident in disciplines such as 'Multidisciplinary Studies', 'Health Sciences', and 'Sciences and Engineering', where the need for concise, impactful titles might be driving researchers to turn to GAI for assistance.

However, this increasing reliance on GAI raises important concerns about the originality and authenticity of academic outputs. While AI tools can enhance linguistic diversity and creativity in title generation, there is a risk that they may contribute to an homogenization of language. The uniform increase in certain verbs across different fields suggests that GAI might be nudging academic titles towards a standardized format, potentially reducing the individuality and innovation that are hallmarks of scholarly knowledge. This trend prompts questions about the role of GAI in academia and how it might shape the future of scientific communication.

As GAI continues to evolve, it is crucial for the academic community to develop standards and guidelines for its use in research writing. Editors, publishers, and reviewers must consider the ethical implications of GAI in academic communication. Guidelines should ensure that while GAI tools can aid in the efficiency and clarity of academic writing, they do not compromise the originality and creativity of scholarly work. Transparency in disclosing the use of AI in the writing process will be key in maintaining academic integrity and trust in scholarly communication.

In conclusion, while GAI offers substantial benefits in terms of productivity and linguistic enhancement, its growing presence in academic writing requires careful scrutiny. The trends observed in this study highlight the dual-edged nature of GAI's influence: it can facilitate innovation and creativity but may also lead to standardization and a loss of individual academic voice. Future research should continue to monitor these linguistic shifts and assess the long-term impact of GAI on academic writing practices. This study provides a foundation for further exploration into the evolving role of AI in shaping the lexicon of scientific communication and underscores the need for proactive measures to ensure that AI's influence enhances, rather than diminishes, the quality and authenticity of academic output.

AUTHOR CONTRIBUTIONS

RC conceived the research, RC and AK developed the methodology, RC, AK and DK did the data gathering or fieldwork, RC, AK and DK wrote and reviewed the article, RC submitted the article.

ACKNOWLEDGEMENTS

Article elaborated within the framework of the project PID2022-141031NB-I00, funded by MICIU/AEI /10.13039/501100011033/ and by ERDF A way of making Europe. RC wishes to thank the Vice-rector of Research at UIB for the grant 15/2024 that permitted the collaboration with the members of West Attika University.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Supplementary material 1 can be accessed at https://figshare.com/s/de04b27e0d548c1af3fc. Bibliometric data available on request to the corresponding author.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article:
Data S1: Supplementary Information.

REFERENCES

Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, *75*, 102370. https://doi.org/10.1016/j.techsoc.2023.102370

Curry, N. (2023). How should digital tools for writing be evaluated?: Reflections from digital pedagogies and applied linguistics. *Journal of Academic Writing*, *13*(1), 53–58. https://doi.org/10.18552/joaw.v13i1.970

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40*(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and

implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fox, C. W., & Burns, C. S. (2015). The relationship between manuscript title structure and success: Editorial decisions and citation performance for an ecological journal. *Ecology and Evolution*, *5*(10), 1970–1980. https://doi.org/10.1002/ece3.1480

Godoy, L. F. (2020). Escritura digital y colaborativa: una práctica discursiva multifacética: Estado del arte y perspectivas para el futuro. *Quintú Quimün*, *4*, 1–29. http://hdl.handle.net/11336/131913

Golan, R., Reddy, R., Muthigi, A., & Ramasamy, R. (2023). Artificial intelligence in academic writing: A paradigm-shifting technological advance. *Nature Reviews Urology*, *20*(6), 327–328. https://doi.org/10.1038/s41585-023-00746-x

Gray, A. (2024). *ChatGPT" contamination": estimating the prevalence of LLMs in the scholarly literature*. arXiv preprint arXiv. 2403.16887.

Haider, J., Söderström, K. R., Ekström, B., & Rödl, M. (2024). GPT-fabricated scientific papers on Google Scholar: Key features, spread, and implications for preempting evidence manipulation. *Harvard Kennedy School (HKS) Misinformation Review*, *5*(5), 1–16. https://doi.org/10.37016/mr-2020-156

Han, Y., Zhao, S., & Ng, L. L. (2021). How technology tools impact writing performance, lexical complexity, and perceived self-regulated learning strategies in EFL academic writing: A comparative study. *Frontiers in Psychology*, *12*, 752793. https://doi.org/10.3389/fpsyg.2021.752793

Kobak, D., Márquez, R. G., Horvát, E. Á., & Lause, J. (2024). Delving into ChatGPT usage in academic writing through excess vocabulary. arXiv preprint arXiv:2406.07016.

Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., … Zou, J. Y. (2024). Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. arXiv preprint arXiv:2403.07183. https://doi.org/10.48550/arXiv.2403.07183

Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. *Perspectives on Medical Education*, *12*(1), 261–270. https://doi.org/10.5334/pme.1072

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570–581. https://doi.org/10.2139/ssrn.4389887

Májovský, M., Černý, M., Kasal, M., Komarc, M., & Netuka, D. (2023). Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *Journal of Medical Internet Research*, *25*, e46924. https://doi.org/10.2196/46924

Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P. C., … Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *Biodata Mining*, *16*(1), 20. https://doi.org/10.1186/s13040-023-00339-9

Sagi, I., & Yechiam, E. (2008). Amusing titles in scientific journals and article citation. *Journal of Information Science*, *34*(5), 680–687. https://doi.org/10.1177/0165551507086261

Samar, R. G., Talebzadeh, H., Kiany, G. R., & Akbari, R. (2014). Moves and steps to sell a paper: A cross-cultural genre analysis of applied linguistics conference abstracts. *Text & Talk*, *34*(6), 759–785. https://doi.org/10.1515/text-2014-0023

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature*, *613*(7945), 620–621. https://doi.org/10.1038/d41586-023-00107-z

Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, *131*, 33–48. https://doi.org/10.1016/j.compedu.2018.12.005

Uribe, S. E., & Maldupa, I. (2024). Estimating the use of ChatGPT in dental research publications. *Journal of Dentistry*, *149*, 105275. https://doi.org/10.1016/j.jdent.2024.105275

Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, *621*(7980), 672–675. https://doi.org/10.1038/d41586-023-02980-0

Zheng, H., & Zhan, H. (2023). ChatGPT in scientific writing: A cautionary tale. *The American Journal of Medicine*, *136*(8), 725–726. https://doi.org/10.1016/j.amjmed.2023.02.011