

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

AI Business Writing Assistant Tools in Practice: Text Mining and Statistical Analysis of User Experience

Yousra Harb, Sérgio Moro & Ayman Harb

To cite this article: Yousra Harb, Sérgio Moro & Ayman Harb (2026) AI Business Writing Assistant Tools in Practice: Text Mining and Statistical Analysis of User Experience, International Journal of Human–Computer Interaction, 42:11, 8193-8222, DOI: 10.1080/10447318.2025.2563743

To link to this article: https://doi.org/10.1080/10447318.2025.2563743

Published online: 30 Sep 2025.

Submit your article to this journal

Article views: 288

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hihc20

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2026, VOL. 42, NO. 11, 8193–8222
https://doi.org/10.1080/10447318.2025.2563743

Taylor & Francis
Taylor & Francis Group

Check for updates

AI Business Writing Assistant Tools in Practice: Text Mining and Statistical Analysis of User Experience

Yousra Harb[a], Sérgio Moro[b,c] and Ayman Harb[c]

[a]Yarmouk University, Irbid, Jordan; [b]Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal; [c]The University of Jordan, Amman, Jordan

ABSTRACT

This study employs a multi-method quantitative research design to review various popular AI writing assistant tools using 11562 actual users' satisfied and dissatisfied reviews. The research design applies Kruskal-Wallis test to differentiate between AI writing assistants based on user rating, functionality, ease of use (EOU), value for money, and customer support attributes; uses the quantitative text analysis to identify satisfaction and dissatisfaction attributes; and employs topic modeling to uncover the specific related attributes that influence user satisfaction. The statistical analysis results indicated that there is a significant difference among AI writing assistant tools according to user rating, functionality, EOU, and customer support attributes. The text analysis revealed the attributes leading to user satisfaction and dissatisfaction. Topic modeling further identified both common and tool-specific attributes. This study provides a relatively comprehensive set of attributes and sub-attributes for evaluating AI writing assistant tools, which can serve as a foundation for future research.

KEYWORDS

AI business writing assistant tools; user reviews; text analysis; user satisfaction

1. Introduction

In today's fast-paced digital landscape, businesses are turning to AI-powered writing assistants to revolutionize how they create, refine, and deliver content (Tunmise & Amanda, 2024). These tools—ranging from grammar and spell-checking to full-text generation—promise to streamline workflows, enhance communication, and boost productivity (Patel et al., 2024; Tung, 2024). Platforms such as ChatGPT, Grammarly, Jasper, and other emerging solutions have become essential allies for businesses seeking to maintain a competitive edge. Yet, despite their transformative potential, not all AI writing tools are created equal. User satisfaction varies widely (Pasch & Ha, 2025), raising critical questions: What makes some tools stand out while others fall short? And how can businesses choose the right AI writing assistant to meet their unique needs?

The rise of AI writing assistants has clearly reshaped the content creation process. From crafting compelling marketing copy to ensuring polished internal communications, these tools offer unparalleled advantages: saving time, ensuring consistency, reducing reliance on human editors, and enhancing the quality of content (Alaofin, 2024; Park et al., 2024; Pattanayak, 2022; Usman et al., 2024). However, the true value of AI tools lies in how well they adapt to diverse business requirements, how intuitive their interfaces are, and how responsive their customer support teams can be (Roslan & Ahmad, 2023; Shrestha et al., 2019). Hence, their success relies on their ability to meet user expectations for quality, accuracy, and ease of use (Garcia, 2025; Lu et al., 2024; Mustață et al., 2025). For businesses, the stakes are undeniably high. Selecting the right tool can determine whether content creation flows smoothly or becomes a source of persistent inefficiencies and frustration (Goh, 2024).

The main value of AI writing assistants is not just their capabilities but also their ability to seamlessly integrate into existing corporate operations. Wamba-Taguimdje et al. (2020) pointed out that the usefulness of these technologies is strongly contingent on how well they integrate with and improve current processes, rather than they act as stand-alone solutions. This highlights how crucial it is to take

CONTACT Yousra Harb yousra.harb@yu.edu.jo Yarmouk University, Irbid, Jordan
© 2025 Taylor & Francis Group, LLC

into account workflow compatibility when deploying AI writing technology in a business environment. For companies, the value of AI writing tools extends beyond simply increasing individual productivity. Hence, AI writing assistant tools should be integrated into necessary digital infrastructures, such as enterprise content management systems, customer relationship management (CRM) platforms, and other vital operational systems (Alexander, 2024; Deep & Zanke, 2024; Ledro et al., 2023). When integration fails, firms face serious risks such as operational inefficiencies, compatibility concerns, and a fragmented user experience. These issues have the potential to reduce the transformational power of AI writing tools (Narayanaswamy, 2024).

While AI marked for its potential in increasing business efficiency, value, and reducing costs, businesses must also look carefully for AI broader implications. More specifically, certain key issues must be considered by businesses when using AI tools such as AI impact on business quality, brand consistency, and information privacy. In this vein, recent research highlights that many companies are worried about data privacy while sharing it with third-party AI services (Culotta & Mattei, 2024). Moreover, AI-generated content in businesses with stringent rules, such as banking, healthcare, and legal services, must conform to strong industry standards and compliance requirements. In these business domains, accuracy and accountability are not optional, but basic needs (Hosanagar & Krishnan, 2024).

These concerns highlight the need for deep evaluation before integrating such tools into business operations. Additionally, the adoption of AI writing assistants in business contexts is still an underexplored area within academic research. Silva and Moro (2021) revealed that while broad AI adoption has been widely studied, its specific application to business writing has gotten less attention. For instance, some prior research studies have investigated the adoption and use of AI writing tools in a business context (e.g., DeJeu, 2024; Zhao & Wang, 2024). However, most have focused on the challenges associated with the adoption of AI writing tools rather than user experience. A related investigation by Retkowsky et al. (2024) sought to understand how employees utilize ChatGPT to perform different tasks, such as structuring content and writing a first draft, using a single methodological approach such as qualitative interview. These gaps highlight the need for further research into the post-adoption experience by exploring how businesses are integrating these tools into their operations and the distinct constraints they confront in doing so. Therefore, our study adopts a multi-method quantitative research design that bridge these gaps by combining statistical analysis with text analysis and topic modeling to effectively address the user experience perspectives of using AI writing assistant tools. This study offers a thorough and detailed assessment of user satisfaction within business environments. Through the integration of quantitative analysis with insights from user reviews, we aim to discover what truly drives satisfaction/dissatisfaction with these tools. Which platforms excel in functionality and ease of use? How does customer support influence user loyalty? And what recurring themes emerge in user feedback? Through this exploration, we provide practical recommendations to help businesses make informed decisions and guide developers in refining their tools for future success.

More specifically, this study is guided by two primary research objectives. First, a statistical analysis (Kruskal-Wallis test) is conducted to evaluate user satisfaction and feature performance across leading AI writing assistants, such as Grammarly, Jasper, Writesonic, and other prominent tools. This analysis compares overall user satisfaction levels to determine which tools are gaining user trust and which are lagging behind. In particular, the study analyzes user ratings to identify the most essential features - such as core functionalities, ease of use (EOU), value for money, and customer support - that significantly influence satisfaction and contribute to building user loyalty. Second, an analysis of user reviews is performed to explore user experiences. This includes quantitative text analysis to identify the most frequently cited features, highlight what users value most and where they perceive shortcomings in AI writing assistants. In addition, the topic modeling algorithm is utilized to extract prevailing themes from user feedback, and to offer insights into the primary pain points and positive aspects that shape user perceptions.

The remainder of the study is organized as follows: Section 2 presents the literature review. Sections 3 details the study methodology. The results are provided in Section 4, followed by discussion in section 5. Finally, conclusion is outlined in section 6.

2. Literature review

2.1. Artificial Intelligence writing assistant tools

Artificial Intelligence (AI) is here to stay, providing numerous advantages for a wide array of distinct daily tasks, from helping in navigation to filter undesired contents such as spam emails (Güngör, 2020). A branch of AI that has gained important traction in recent years is generative AI, in which AI is used not only to identify patterns of knowledge and make predictions, but also to generate new meaningful content for a given problem (Sætra, 2023). Unarguably a specific area that recently had a staggering growth in the last few years is human language, where generative AI is having a disruptive impact in most tasks that involve natural language processing (NLP), such as reading, interpreting human textual contents, and writing a reply in a knowledgeable interactive text that reflects some degree of awareness about the context and problem at hands (Barreto et al., 2023). This is mostly because of the development of Large Language Models (LLM), that are trained with corpora of billions of tokens to develop deep learning neural networks that can "understand" how words and sentences relate to each other (Raiaan et al., 2024).

The computational capabilities currently available have provided accurate generic language models that can mimic to a certain degree a human actor. However, there are only a handful of very large organizations that can have access to huge computational resources to train such models (Hosanagar & Krishnan, 2024). These include publicly available models, such as ChatGPT by OpenAI, which has become the mainstream tool widely adopted by individuals worldwide, or the Copilot by Microsoft, adopted by many organizations spread through the globe that already have licenses for Microsoft suite products (Narayanaswamy, 2024).

As a result of this paradigm change, scholars turned their attention to how society is incorporating these tools while at the same time changing individual and group behaviors in face of daily problems that now can be eased by typing on a keyboard for assistance (George & George, 2023). Scholars recognize the impact of generative AI in how organizations operate, although there is no consensus on the best path to follow to optimize the use of such tools (Giordano et al., 2024).

2.2. Business writing assistant tools

We conducted a systematic literature review to assess current state of the art in business writing assistant tools. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology was chosen given its popularity for conducting systematic reviews across different fields (Page & Moher, 2017). It defines a sequence of steps consisting of (1) identifying relevant sources, collecting them, and preparing for subsequent analysis, (2) screening the collected documents, (3) selecting those suitable for the topic being studied, and (4) conducting a synthesis on the final set of articles. Figure 1 shows the results of each undertaken step.

The Scopus database was adopted, considering it is one of the most widely used databases for querying scientific literature. By indexing articles from reputable peer-review journals, it has consistently been chosen for literature analyses in studies in a wide range of fields, such as education (Singh & Miah, 2020) and marketing (Silva & Moro, 2021). Scopus enables to build complex queries by combining articles metadata (e.g., source type – journal, conference proceeding; language; sections – title, abstract, keywords) with content search using Boolean operators.

The query used for retrieving the body of literature includes journal articles or conference proceedings articles, since these two are the most significantly recognized outlets for publishing novel scientific contributions, whereas only English written literature was included, considering that the international academic community recognizes English as the common denominator language for communicating internationally with peers. Additionally, preliminary tests revealed that there was a large portion of the retrieved articles were related to chat bots and similar tools, thus we chose to exclude them, as our goal was to include only writing assistant tools. Considering also the popularity of some tools such as ChatGPT and Copilot, we decided to include these directly in the query along with the list of writing tools identified by the Gartner.com study (2024), namely, Grammarly, Jasper, Writer, Writesonic, WordAi, TextWizard, Ginger Business, ProWritingAid. The final query used was as follows:

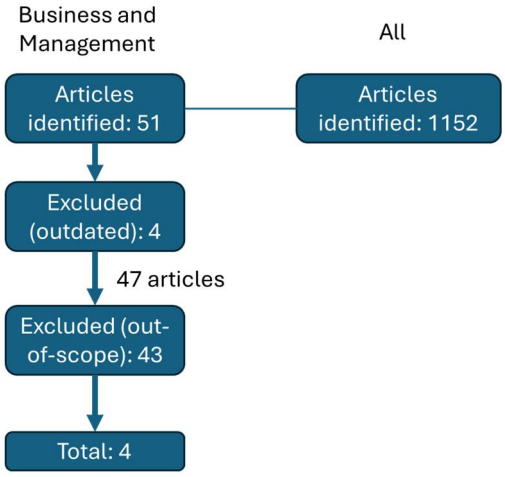


Figure 1. PRISMA flow adopted.

TITLE-ABS-KEY((business OR management) AND writing AND ("assistant tool" OR GPT OR ChatGPT OR Copilot OR Grammarly OR Jasper OR Writer OR Writesonic OR WordAi OR TextWizard OR "Ginger Business" OR ProWritingAid) AND (AI OR "artificial intelligence" OR LLM OR "language model") AND NOT (ChatBot OR chat)) AND (LIMIT-TO (SRCTYPE,"j") OR LIMIT-TO (SRCTYPE,"p")) AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"cp")) AND (LIMIT-TO (LANGUAGE,"English"))

The result was a total of 51 articles as of January 2025. We then hypothesized that these tools while popularized in recent years, its use and consequently its research adoption and impact has been quite limited within business and management activities in comparison to other areas such as education. Therefore, we searched Scopus using the same query but removed the "(business AND management)" condition. The result was a staggering total of 1152 articles, corroborating our assumption, and clearly showing lack of research within business and management context.

Of the 51 articles, four were based on outdated tools (1994; 1990; 1986; 1985) and were therefore excluded, leaving a final total of 47 articles. Figure 2 shows the distribution over the years for all categories (blue bars, without the "business and management" filter) and after filtering to include only "business and management" (orange bars). As expected, given the topic is relatively new, the number of studies has boomed after 2020. However, that increase is particularly steep for "business and management," since the years of 2020, 2021, and 2022 saw only one published article per year, while after 2023 the increase was exponential, with current year of 2025 already resulting in three published articles, suggesting such trend will not fade in the following years. By comparing both types of bars, we emphasize that research specifically related to business writing assistant tools still lags behind other domains, justifying the research we present in this study.

Considering the present study aims at comparing AI writing assistant tools for business, we scrutinized the 47 previously selected articles by reading each one's title, abstract and keywords to assess their relevance to our proposed contribution. Specifically, we aimed at selecting only articles that could be directly framed under a business/management context. Both words "business" and "management" can be used in many contexts, outside the scope of enterprises and organizations, such as using AI writing tools to assist in academic reports/assignments in business and management schools (e.g., Zhan et al., 2023). Another example was the study by Phillips and Robie (2024), which aimed at understanding the challenges raised by AI tools in human resources assessment. Another large fraction of those were focused on educational contexts, with several related to medicine education and research (e.g., Zumsteg & Junn, 2023). We found also several studies focusing on the problem of distinguishing between human writer and AI writer and its ethical and legal implications (e.g., Aburass & Abu Rumman, 2024). Thus, the remaining four studies that are focused on business/management writing assistance are summarized in Table 1.

There are a few key takeaways from Table 1 that help in positioning the contribution of the present study. Both the studies by Zhao and Wang (2024) and by Desai et al. (2024) aimed at studying the use of ChatGPT in the context of accounting and human resource management, respectively. Moreover, their

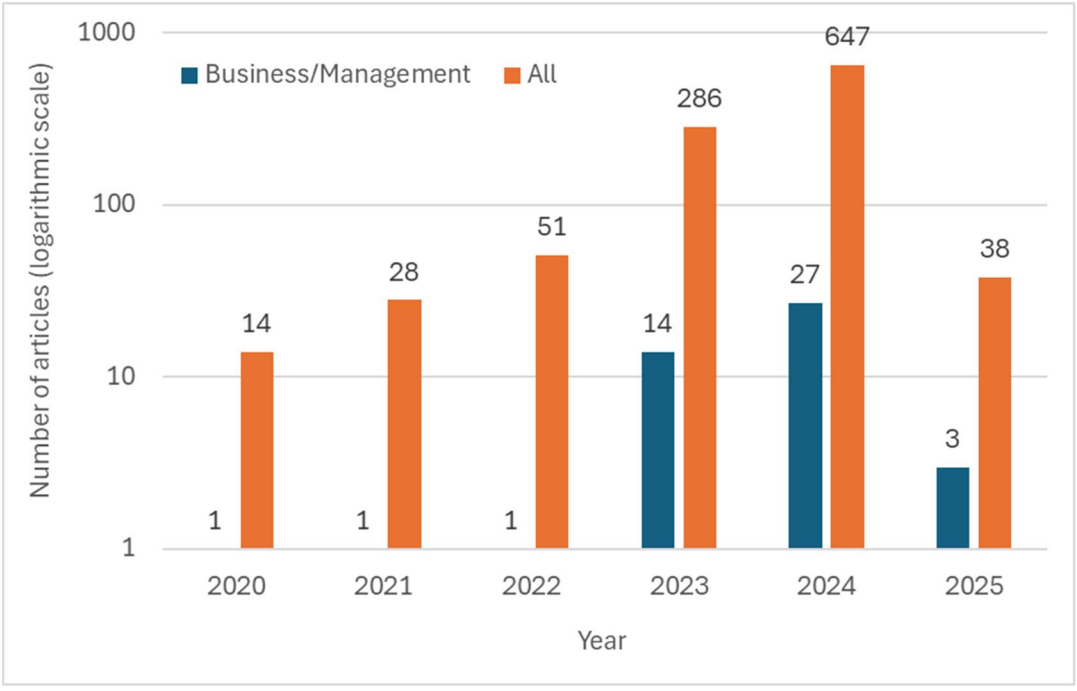


Figure 2. Distribution of articles over the years (after 2020).

Table 1. Business writing assistant tools.

Reference	Tool	Summary
(Zhao & Wang, 2024)	ChatGPT	Explored applications and challenges of ChatGPT in accounting, including writing accounting reports (as well as other accounting tasks)
(Desai et al., 2024)	ChatGPT	Focused on the specific problem of gender bias in writing recommendation letters for academic surgery positions
(DeJeu, 2024)	n.a.	Focused on adapting business writing instruction for the AI age, using real-world business examples
(Retkowsky et al., 2024)	ChatGPT	Focused on the problem of understanding how employees use ChatGPT for different tasks, including: structuring content; writing a first draft; embellishing text; and proofing work

focus was on the challenges that the use of ChatGPT on those contexts bring to the companies. Thus, the user experience of adopting ChatGPT was not assessed at all. On another different perspective, DeJeu (2024) addressed the challenges of instructing AI tools to write text in a business context, without dealing with the employees' feedback on their experience when using such tools. Finally, Retkowsky et al. (2024) studied the impact of managing employees by understanding how they use ChatGPT.

Thus, from all the four studies in Table 1, none effectively addresses the user experience perspectives of using AI writing assistant tools, setting the research gap addressed by our study. Accordingly, we raise the following research questions:

RQ1. To what extent business AI writing assistant tools differ significantly based on users' overall rating, EOU, functionality, value for money, and customer support attributes?

RQ2. What business AI writing assistant tools attributes that are leading to user satisfaction and dissatisfaction?

RQ3. What the attributes or features that are considered specific to business AI writing assistant tools?

3. Methodology

3.1. Data collection

Users' reviews, e.g., satisfied and dissatisfied are used as a source of feedback on AI writing assistant tools for business. In fact, user reviews are an essential source for highlighting the key product aspects

and features of greatest concern (Yang & Fang, 2004). Thus, an analysis of users' reviews including the pros and cons comments would be useful to explore the rich aspects of the AI writing assistant tools such as functionality, value for money, ease of use (EOU), and customer support. We accessed the users' online reviews from SoftwareAdvice website in July 2024, a consulting and review platform that allows experienced users with using software products to share their complaints and compliments. We decided to gather users' reviews on AI writing assistants from SoftwareAdvice website as it includes large samples of reviews. Users are also allowed to post both satisfied and dissatisfied reviews as well as provide their rating scores on different aspects of software product. Additionally, this website is utilized in previous literature to gather users' reviews on different software products (e.g., Chowdhury et al., 2021; Harb et al., 2024).

The data collected from SoftwareAdvice includes textual data and criteria ratings expressed by users. Specifically, when users evaluate AI writing tools, they typically assess them from various dimensions. The user's evaluation includes textual reviews and criteria ratings score of different aspects of the tool. The textual data includes the users' reviews and categorized into "PROS" which reflects satisfactory reviews and "CONS" which describes user's dissatisfactory comment. The criteria ratings score with values ranging from 1 to 5, on the other hand, comprise numerical data on the following attributes: overall rating, functionality, value for money, EOU, and customer support. Below is a sample review selected from SoftwareAdvice. Data preprocessing procedures are utilized on textual data as explained in section 3.2.3.

Overall rating: 3/5
Functionality: 3/5
Value for money: 3/5
Ease of use: 4/5
Customer support: 3/5
***PROS**: It is incredibly easy to setup within the organization and use. Most of the spelling and punctuation errors are covered by it.*
***CONS**: Sometimes suggestions are misleading and wrong. Confusion between the languages can still happen and suggest things that no one wants changed (places or names). Some style suggestions are questionable. Limited to English only and thus not flexible enough.*

In this study, we accessed all reviews, by July 2024, of textual data and users' ratings score on the following AI writing assistant tools: Grammarly, Jasper, Writer, Writesonic, WordAi, TextWizard, Ginger Business, and ProWritingAid. We selected the tools for review based on the Gartner.com (2024) website's classification of these tools as top performers in AI writing assistants based on recent software buyers ranking. However, we excluded any tools that had fewer than 50 reviews from further analysis to minimize the impact of incomplete data, as not all users provide detailed textual feedback (satisfied/dissatisfied comments) or comprehensive rating scores on the different aspects of the tools. The total number of online reviews gathered from this website is 11562. Figure 3 presents the sample reviews distribution for the AI writing assistant tools.

3.2. Data analysis

This study employs a quantitative multi-method research design. This approach combines statistical analysis of numerical data, quantitative text analysis using frequency analysis, and text mining using topic modeling of users textual reviews. The usage of this methodological approach involves using different analytical methods to analyze data to answer the research questions. Multi-method research design increases the chance of identifying varied and extensive findings and enhancing the overall depth of the study (Kasirye, 2021). The following three subsections detail the quantitative multi-method components of this study.

3.2.1. Statistical analysis of online reviews quantitative ratings

Following data collection, two types of data formats were collected: quantitative ratings (1–5, as previously explained), and textual comments. In a first stage, a statistical analysis was conducted using the quantitative ratings. Therefore, to answer the first research question that concerns assessing user

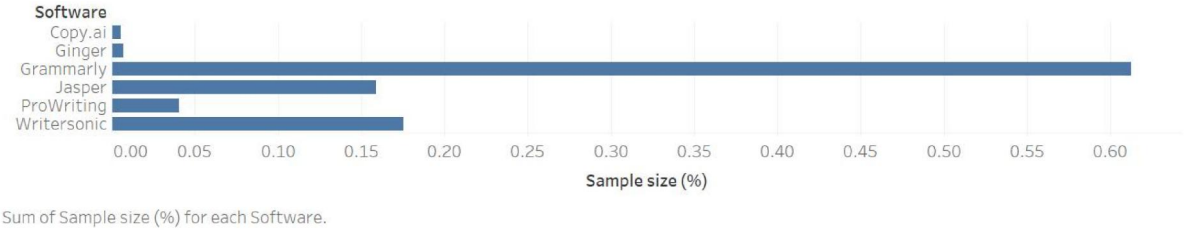


Figure 3. Sample reviews distribution.

satisfaction with different AI writing assistant tools for business across various attributes including user overall ratings, functionality, value for money, EOU, and customer support. The research method utilizes a statistical analysis to assess whether there are significant differences in users' ratings, who have experienced these tools in their organizations, across AI writing assistant tools. In this study, we used Kruskal-Wallis test to compare the average ranks of five attributes (user rating, EOU, functionality, value for money, and customer support) across different AI writing assistant tools.

3.2.2. Categorization of online reviews text

To answer the second research question, which concerns identifying which AI writing assistant tools' attributes are leading to user satisfaction in the users' positive reviews and which are leading to user dissatisfaction in the negative reviews of AI writing assistant tools, a textual analysis is required. The key is to identify the relevant words within the reviews that can be related to the different categories to be analyzed. In the first step, we extracted specific attributes of AI writing assistants associated with user overall ratings (satisfaction), functionality, EOU, value for money, and customer support. These attributes were integrated into the users' online reviews dataset and users provided their rating scores for each attribute based on their usage experience and satisfaction. Given the large volume of the reviews, it makes sense to extract these predefined attributes from the textual content of the users' online reviews. In order to keep the scope within these attributes in the users' reviews, we created a dictionary that includes a list of categories, keywords or items linked to the predefined attributes, and the theoretical grounding frameworks (Table 2).

This dictionary is developed based on a deductive approach (Elias, 2012). Specifically, we mapped each predefined AI writing assistant tools' attribute to relevant theoretical constructs drawn from well-established theoretical frameworks (see Table 2). These established models informed the operationalization of each attribute through reviewing empirical studies that applied these models to examine how these studies operationalized the constructs in their survey instruments or conceptual definitions. Thus, we then conducted extensive review of existing empirical literature that employed these theoretical frameworks to examine how these attributes were identified and operationalized. The theoretical models served as theoretical lens to guide the tools' key attributes operationalization. This enabled a structured identification and quantification of users' satisfaction dimensions and their associated indicators in the dataset. This process ensures that the development of the dictionary attributes is theoretically justified and empirically validated.

Specifically, user overall rating reflects the overall user satisfaction which refers to how pleased or content an individual is with a technology (Oliver, 1980). In our study context, user satisfaction refers to the extent to which AI writing assistants' users satisfy, happy, or convinced with the tool's functionality, EOU, value for money, and customer support. Thus, these factors contribute to the overall user satisfaction and the users' ratings of these attributes are influenced by their satisfaction levels. Therefore, we evaluated user satisfaction with AI writing assistants by analyzing it through the four attributes: functionality, EOU, value for money, and customer support. Previous literature has been found that these attributes are predictors of user satisfaction (e.g., Anugrah, 2020; Hayajneh & Harb, 2023; Rajaguru, 2016). We defined these attributes based on well-established theoretical frameworks such as technology acceptance model (TAM) (Davis, 1989), the expectation confirmation model (ECM) (Oliver, 1980), and the information systems success model (D & M model) (DeLone & McLean, 2016). These theoretical frameworks have been extensively validated in the literature. Specifically, based on TAM, we defined the

Table 2. The attributes dictionary.

Attribute	Category	Keywords	Theoretical framework
Functionality	System/software quality	adaptable, available, reliable, usable, integration, reporting, accessible, flexible, functionality	D & M model (2016)
	Performance	performance, efficiency, time-saving/waste less time, quality of work, accuracy/accurate, powerful, completeness, response time	
Value for money	Value for money	good value, worth the price/worth the money, cost-effective, affordable, economical, good price/fair price/reasonable price, pricey, expensive	VAM (2007)
Customer support	Responsiveness	responsiveness, prompt response, prompt service, on-time response, quick response, responsive communication	SERVQUAL (1988) and e-SERVQUAL (2001)
	Reliability	reliability, software failure, software shutdown, error-free	
	Competence/Assurance	competence, assurance, answer questions, service support, customer support, solve problem, assistance, troubleshooting, resolution, technical support, professional support	
	Security	security, privacy, protection, lock out from service	
	Empathy/Relationship service	information availability, information content, customization, personalization, online help/tutorials, training, support material, user instruction, continuous improvement, update speed	
Ease of use	Understandability	understandable, readability, effortless, clear, learning curve, quick to learn, easy to learn, easy to master	TAM (1989)
	Ease of use	easy to use/easily use, easy to navigate, easy to search	
	Friendly	user-friendly, intuitive, simple, straightforward	

associated items with EOU, which refers to the extent to which an individual believes that using a particular system would be effortless and easy to learn (Davis, 1989), we referred to TAM and drew from a wide spectrum of relevant empirical research studies that have employed TAM to measure the EOU such as (Ashfaq et al., 2020; Hong et al., 2006; Rafique et al., 2020).

Functionality attribute, on the other hand, is defined as a set of attributes that enable the user to perform different activities (Elias & Lubua, 2024). In the information system domain, it conceptualized as the quality of the system or application to deliver intended work through input, process, and output (Peters et al., 2016). Functionality is also linked to system usability for interacting and modeling with the data (Peters et al., 2016) as well as system quality (e.g., Jo & Bang, 2023). Functionality attribute items set were adapted from the following literature that directly studied functionality feature such as (Peters et al., 2016) and from research studies that employed D & M model to study the concept of system quality (e.g., Hayajneh & Harb, 2023; Al-Okaily, Teoh, & Al-Okaily, 2023; Nelson et al., 2005).

Further, in utilitarian perspective, value for money, describes the perceived utility in relation to the price paid (Jayanti & Ghosh, 1996). It measures what the customer has received compared to what has been given (Rajaguru, 2016). This attribute's item set was adapted based on the Value-Based Adoption Model (VAM) (Kim et al., 2007) as well as the empirical research studies such as (Alalwan et al., 2017; Rajaguru, 2016).

Customer support attribute, on the other hand, is conceptualized in the literature as service quality and refers to the extent to which the user of a system or an application receives the expected level of services provided by service provider (DeLone & McLean, 2003). SERVQUAL model (Parasuraman et al., 1988) is among the major models employed in the existing literature to measure the traditional service quality dimensions. The measurement items used in this model to define service quality are grouped under five dimensions: tangible, reliability, responsiveness, assurance, and empathy. However, this model was extended to the domain of information systems context (DeLone & McLean, 2003). For instance, some research studies (e.g., Zeithaml et al., 2000) built upon the SERVQUAL to develop e-SERVQUAL to define the appropriate dimensions for measuring service quality in the online context such as efficiency, reliability, fulfillment, privacy, responsiveness, contact, etc. Therefore, in the current study, we drew upon SERVQUAL and e-SERVQUAL, to define the items set to measure the service quality attribute in the context of AI writing assistant tools.

To extend the validity of this dictionary, we also examined relevant literature that identified these attributes based on actual users' reviews and how users in their feedbacks describes these attributes. For

instance, Yang and Fang (2004), in their study, employed content analysis approach of 740 customers reviews to define 16 major online quality dimensions and their associations with customer satisfaction. It is worth mentioning that the items used to define the attributes in our study, either derived from the existing relevant literature or from the actual online users' reviews studies, are contextually adapted to suit the study context. This dual approach of creating the dictionary provides a comprehensive approach on how AI writing assistants attributes are identified in theoretical frameworks found in academic research, complementing the practical interpretations of these attributes from the perspective of actual users. In the second step, we developed a Python script to perform textual analysis, specifically designed to calculate the frequency of each term for each attribute within the given dictionary.

3.2.3. Text mining

The third research question concerns identifying additional features or attributes specific to AI writing assistants beyond those identified in the pre-defined AI writing assistants' attributes dictionary. To answer this question, a Latent Dirichlet Allocation (LDA) topic-modeling algorithm is implemented to uncover the features the are specific to AI writing assistants from the online reviews posted by actual users. The process of driving the AI writing assistants' features includes the following steps: (1) Online reviews preprocessing. This step includes the removal of stop words. We also expanded the list of the stop words to remove the noisy words such as ("every," "also," "another," "although," "might," "according," etc) and to enhance the quality of the extracted topics. Additionally, to ensure more focused topics and precisely identify the specific attributes that tied to AI writing assistant tools, we removed keywords such as the name of the tools, "easy," "use," as we observed that these keywords frequently happened in the returned topics. The next step in the reviews preprocessing involves converting all words to the lowercase, and performing lemmatization to reduce the words to the base form. (2) LDA algorithm as one of the most important generative probabilistic used algorithm was employed to discover the topics from the textual documents (Blei et al., 2003) using Python software. In this unsupervised generative process, each document is represented as a multinomial distribution across K topics, with each topic being modeled as a multinomial distribution over the vocabulary V. One of the key parameter of LDA is determining the number of topics (k). To choose the right number of topics, this study used the coherence measure (Wallach et al., 2009). Specifically, the optimal LDA model was selected based on the highest coherence score. As shown in Figure 4, the number of topics in the positive dataset is $k = 6$ and 36 in the negative reviews dataset. The final model is then trained using the selected number of topics k. (3) Assigning representative names or labels for the generated topics is usually a manual process, particularly when the labeling process requires domain expertise to ensure high-quality results (Harb & Shang, 2022). In the current study, the generated topics of the positive and negative reviews were manually labeled by two of the authors who are knowledgeable in information system domain. The labeling process was performed based on the association of the words in each topic. The labeling was also performed independently by the authors to avoid biases and ensure the

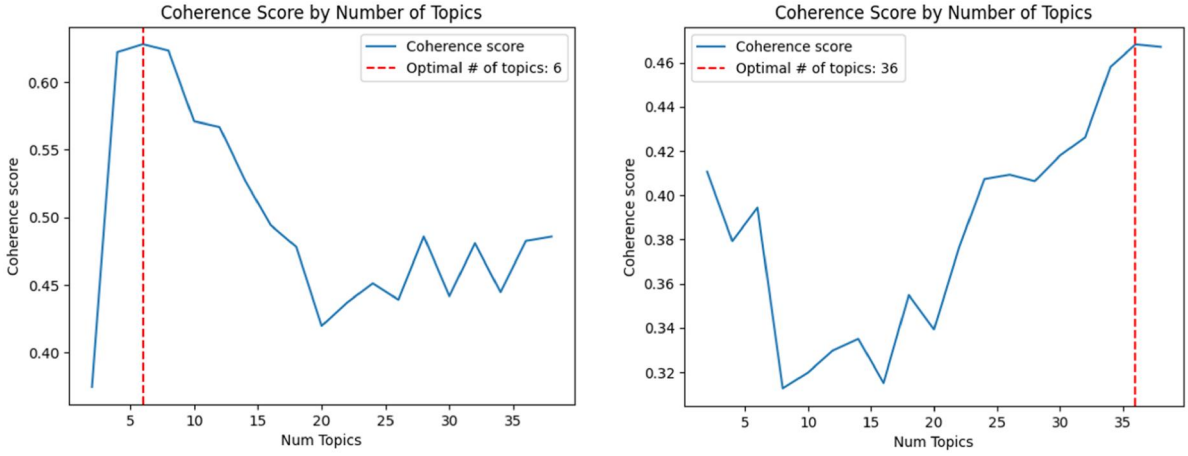


Figure 4. The optimal number of topics in the positive reviews and negative reviews.

quality of the labeling process. Then we used Kappa coefficient to evaluate the level of agreement between the two authors. We found a Kappa score is (0.778, $p = 0.000$), indicating a substantial level of agreement in the labeling process results (Landis & Koch, 1977). Table A1 (Appendix) presents the generated topics from both types of reviews and the topics representative labels.

4. Results

We present the findings corresponding to the three research questions as follows:

4.1. Statistical analysis of online reviews quantitative ratings

The first research question concerns evaluating users' satisfaction with different AI writing assistants for business. Prior to selecting the appropriate statistical test to answer this question, the following assumptions for one-way ANOVA were assessed. These assumptions include independence of observations, homogeneity of variances, and normality (Lix et al., 1996). Independence assumption was evaluated based on the study design (Lix et al., 1996). In the data sample used in the current study, each user provided a single evaluation of the tool used in their organization. Users also represented different companies and roles, with no duplicated responses. Therefore, there was no strong evidence of dependence among observations. Residuals plots were also examined to assess this assumption. The plots of residuals appear to be scattered randomly without forming any systematic patterns, indicating no strong evidence of violation of independence assumption. However, the homogeneity of variances assumption (Parra-Frutos, 2013) was violated as evidenced by Levene test ($p < 0.05$). The normality test was also violated as evidenced by Shapiro-Wilk test ($p < 0.05$) (Shapiro & Wilk, 1965). Given the violation of normality and homogeneity assumptions, an appropriate non-parametric alternative, the Kruskal-Wallis test (Kruskal & Wallis, 1952), to one-way ANOVA was employed (Lix et al., 1996).

Table 3 summarizes the results of Kruskal-Wallis test. As shown in Table 3, there are significant differences in users' ratings across AI writing assistant tools for each of the five attributes (user rating, EOU, functionality, value for money, and customer support). The significance value for these comparisons was 0.000 ($p < 0.05$). However, Kruskal-Wallis does not indicate which tools differ from one another. Therefore, we conducted post hoc analyses to further investigate the differences revealed by Kruskal-Wallis test. Specifically, we employed the Dunn with Bonferroni correction (1964) test, which is a common post hoc test of Kruskal-Wallis and suitable for situations where sample sizes are not equal across groups (Elliott & Hynan, 2011).

The Dunn test revealed significant differences between various AI writing assistant tools for each of the five attributes. Table A2-Appendix presents the pairwise comparisons of each individual AI writing tool with the other remaining tools along with their associated p-values. Tools pairs with (p-values $<=0.05$) are considered to differ significantly in user rating scores.

The Dunn post hoc test results indicated that several pairs of AI writing assistant tools showed significant differences in user ratings. By examining the mean ranks from Kruskal-Wallis output (Table A3-Appendix), we can determine which tools received higher or lower rating scores. The AI writing assistant tools that are statistically different in user overall rating attribute are: (1) Jasper tool (mean rank = 6074.67) which received significantly higher ratings compared to Grammarly ($p = 0.000$), Ginger ($p = 0.000$), and ProWritingAid tools ($p = 0.000$), indicating its high level of user satisfaction. (2) Grammarly (mean rank = 5736.04) and Writesonic (mean rank = 5930.94) tools also received significantly higher ratings compared to Ginger and ProWritingAid ($p = 0.000$ for each), (3) Ginger, On

Table 3. Kruskal-Wallis test results.

Test statistics[a,b]

	Overall rating	EOU	Functionality	VOM	Customer support
Chi-Square	159.497	188.167	183.068	50.717	245.342
df	5	5	5	5	5
Asymp. Sig.	0.000	0.000	0.000	0.000	0.000

a. Kruskal Wallis Test.
b. Grouping Variable: AI Writing Assistant Tools.

the other hand, generally exhibited significant lower ratings relative to the other tools (mean rank = 3310.47; $p = 0.000$).

Further, the Dunn post hoc test results also identified several statistically significant differences in ease of use (EOU) attribute ratings among different AI writing assistant tools: (1) Grammarly (mean rank = 5866.93) received significantly higher ratings compared to Jasper (mean rank = 5563.65; $p = 0.000$), Ginger (mean rank = 4088.72; $p = 0.000$), and ProWritingAid (mean rank = 4730.22; $p = 0.000$). (2) Writesonic (mean rank = 5999.22) also demonstrated significant higher ratings than Jasper, Ginger, and ProWritingAid ($p = 0.000$ for each). (3) Jasper (mean rank = 5563.65) was also rated significantly higher than Ginger, and ProWritingAid ($p = 0.000$ for each). This means Grammarly and Writesonic outperformed in ease of use compared to most of the tools, and Jasper following behind. (3) Copy.ai performed better than Ginger in this attribute ($p = 0.017$). (4) The results also indicated that Ginger and ProWritingAid exhibited the lowest rating in terms of EOU compared to most of the tools.

The post hoc results also showed that there are statistically significant differences among the following AI writing assistant tools in functionality: (1) Jasper tool (mean rank= 6147.65) received the highest rating in terms of functionality compared to Grammarly (mean rank= 5649.66; $p = 0.000$), Copy.ai (mean rank= 4199.69; $p = 0.013$), Ginger (mean rank= 3642.15; $p = 0.000$), and ProWritingAid (mean rank = 5226.42; $p = 0.006$). (2) Writesonic (mean rank= 6151.68) was also rated significantly higher in functionality attribute compared to Grammarly, Copy.ai, Ginger, and ProWritingAid ($p = 0.000$ for each). These findings highlight that Jasper and Writesonic tools are leading in terms of functionality compared to most of the tools. However, similar to the previous results in overall user ratings and EOU attributes, (3) Ginger (mean rank= 3642.15) was rated lower in functionality compared to Grammarly, Jasper, Writesonic, and ProWritingAid ($p = 0.000$ for each).

Furthermore, in term of the value for money attribute, there are statistically significant differences among the following AI writing assistant tools: (1) Grammarly (mean rank = 4428.25) and Writesonic (mean rank= 4658.75) received higher ratings than Ginger (p-value= 0.000 for each). (2) Ginger, on the other hand, had the lowest mean rank in terms of value for money compared to Grammarly, Jasper, Writesonic, and ProWritingAid ($p = 0.000$ for each). This insight is essential for both users seeking cost-effective tools and for developers aiming to improve the perceived value of their tools. (3) Copy.ai (mean rank= 3672.39) did not differ significantly from the majority of the tools (all $p > 0.05$).

The Dunn post hoc test results in customer support attribute revealed that: (1) Jasper tool (mean rank = 4822.79) was rated highest in this measure. It received significantly higher ratings than Grammarly, Writesonic, Ginger, and ProWritingAid ($p = 0.000$ for each). This may indicate that Jasper is perceived as providing high level of customer support compared to most of the tools. (2) Writesonic (mean rank= 4446.17) was significantly rated higher than Grammarly, Ginger, and ProWritingAid ($p = 0.000$ for each). (3) Grammarly (mean rank= 4015.08) was significantly rated higher than Ginger ($p = 0.000$), ProWritingAid (mean rank= 3827.85; $p = 0.000$) and Copy.ai (mean rank= 4026; $p = 0.035$) also rated higher than Ginger in this attribute. This could suggest that Ginger is viewed as offering comparatively less satisfactory support, whereas Jasper is considered to provide the best customer support.

4.2. Identified categories from online reviews text

The second research question concerns identifying which attributes of AI writing assistants are mainly leading to user satisfaction and which are leading to user dissatisfaction. In terms of overall frequency of mention, the most frequently mentioned AI writing assistants' attributes that are leading to user satisfaction, as shown in Figure 5, are: ease of use which was the major factor leading to AI writing assistant's adoption and satisfaction followed by software functionality and then customer support. Conversely, value for money was the most cited as a concern, in the negative reviews, among the four evaluated attributes. The price attribute was a common theme in the negative reviews and centered on the high cost of this technology, which acted as a limitation of user satisfaction.

In terms of attributes category frequency of mention, the most frequently cited categories leading to customer satisfaction, as shown in Figure 6, are the ease of use and user-friendly aspects of the ease of

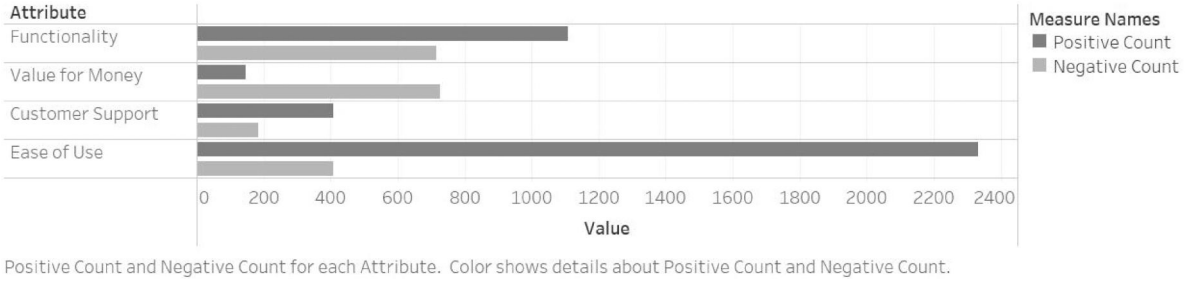


Figure 5. AI writing assistants' attributes frequency.

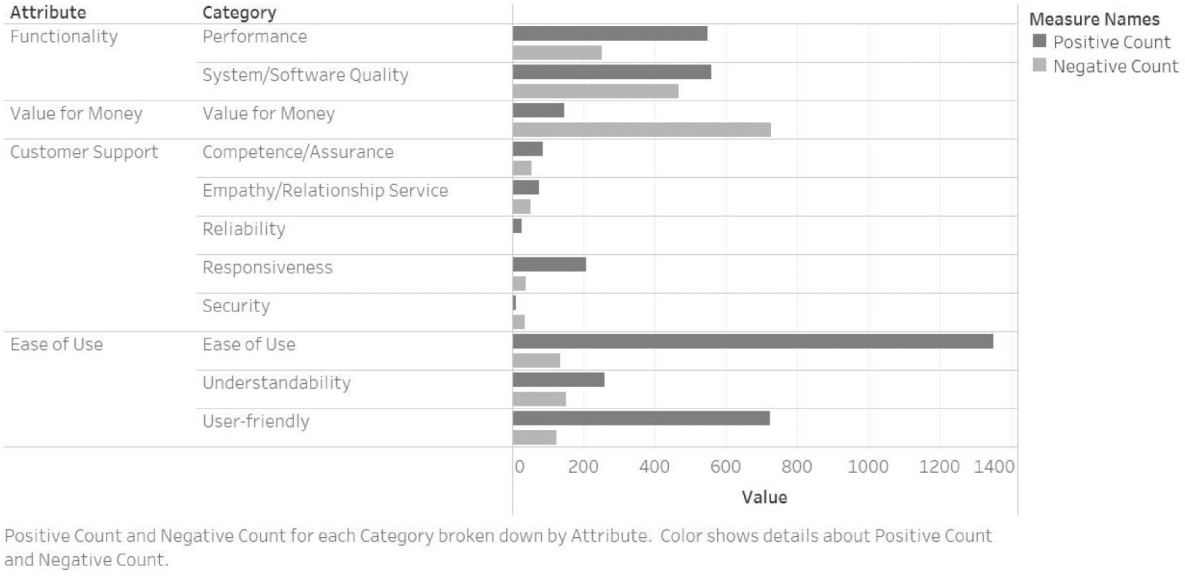


Figure 6. AI writing assistants' attributes categories frequency.

use attribute, followed by the performance and software quality categories of the functionality attribute, and then by responsiveness category of customer support.

Numerous users highlighted the ease of use feature, in the positive reviews, as the best feature of the software, consistently praising it as the key strength of the software. For examples: "The ease of use it impressed me the most." "Best UI and ease of use. Very intuitive menus. Amazing results."

On the other hand, the inspection of the negative reviews, as shown in Figure 6, revealed relatively little concern regarding ease of use attribute and its related categories, as indicated by their low frequency of mention compared to that in the positive reviews. Additionally, the analysis of the negative reviews showed that several users started by valuing and appreciating the ease of use aspects of the software before expressing their concerns about other aspects of the software. This indicates that AI writing assistants' users blend the positive feedback with their concerns. Example from the negative reviews include: " … despite its ease of use, it does necessitate some effort when it comes to installing it on one's device which may end up to be a complex activity to some people who are not that familiar with software."

Although ease of use and its related categories were not frequently raised in the negative reviews, Figure 6 shows that understandability was among the most frequently mentioned. When linked to Table 4, the keyword "learning curve" was the most cited term, which is associated with understandability category, by users in the negative reviews. When analyzing the users' reviews, we found that users have raised the initial learning curve issue when adapting to the tool. Once they overcome it, the software proved to be highly beneficial. " … .I think there is a learning curve to making sure it is turned on in all the applications you use … .," "Additionally, despite the intuitive design, there is still a bit of a learning curve to fully leverage all the features and tools available."

The figure also shows that performance and software quality categories of functionality attribute are comparable in terms of frequency of mentions in the positive reviews. However, software quality was the most cited category in the negative reviews. The further examination of associated keywords frequency in Table 4 reveals that accuracy/accurate of performance category was the most cited by users in the positive reviews. There was a significant appreciation by users about the capability of the software in generating accurate content. "Complex spelling gets easier, while preparing a proposal or website content its accuracy proves the best." This indicates the capability of the tool in improving business writing and detecting errors "It is a helpful writing tool that proofread with great accuracy while helping you decide the tone of your message … ."

Table 4. Identified AI writing assistants' attributes' items frequency.

Attribute1	Category1	Keyword1	Positive count	Negative count
Functionality	Performance	accuracy/accurate	381	189
		completeness	1	1
		efficiency	32	6
		performance	39	22
		powerful	73	30
		quality of work	5	0
		response time	2	2
		time-saving	16	0
	System/Software Quality	accessible	19	26
		adaptability/adaptable	8	4
		availability/available	169	190
		flexibility/flexible	43	11
		functionality/functional	56	71
		integration	165	115
		reliability/reliable	57	28
		reporting	11	5
		usability/usable	31	16
Value for Money	Value for Money	affordable	65	54
		cost-effective	4	2
		economical	1	3
		expensive	28	479
		fair price	2	0
		good price	3	1
		good value	2	3
		pricey	4	125
		reasonable	21	19
		value for money	8	8
		worth the money	6	16
		worth the price	3	16
Customer Support	Competence/Assurance	answer questions	1	0
		assistance	54	16
		assurance	5	1
		competence	1	0
		customer support	24	36
		resolution	1	0
	Empathy/Relationship Service	customization	23	16
		personalization	1	2
		training	44	30
		tutorial	8	4
	Reliability	error free	21	0
		reliability	5	4
	Responsiveness	communication	185	29
		quick response	1	1
		responsiveness/responsive	22	8
	Security	privacy	2	18
		security	10	18
Ease of Use	Ease of Use	ease of use/easy to use	1,314	131
		easy to navigate	36	3
	Understandability	clear	111	37
		easy to learn	14	2
		effortless	15	1
		learning curve	22	71
		quick to learn	1	0
		readability/readable	78	16
		understandable	18	23
	User-friendly	intuitive	170	30
		simple	388	74
		straightforward	49	7
		user-friendly	117	14

On the other hand, the most frequently mentioned keywords in the performance category of functionality in the negative reviews was also accuracy/accurate. Although there was a significant acknowledgment that the software is mostly accurate in the positive reviews, some users had some concerns about the accuracy and reliability of the content generated by the software. In some cases, AI writing assistants produced incorrect or misleading information, which required manual correction as noted by some users "There are times I have to manually edit some of the copies written by Copy.ai because they are not entirely accurate." "It can sometimes generate content that is not factually accurate or that is not in line with your brand voice."

For software quality category of functionality attribute, availability/available and integration were the most frequently mentioned keywords in the positive reviews. The users acknowledged the availability of the software all the time " … i would say its more pricey but you have to expect to pay well for a business product which is available 24/7 to edit and check articles or important documents."

They also valued the availability of various features of the software "I found this tool one of the easiest to get started and the templates available are fantastic for 90% of everyone's needs." "There are few tools on the market that can compete with the variety and options available in this tools."

On the other hand, the same keywords availability/available and integration emerged as the most cited by users in the negative reviews of the software quality category of functionality, as shown in Table 4. This dual pattern suggests that these features play a critical role in shaping user satisfaction, acting as both a strength and point to divergent of user expectations depending on user experience. For instance, users have expressed their concerns about the following aspects: the software only support English and not available in other languages "The only feature I find of this software is that it is only available to correct grammar in English. And is not available for other languages." lack of the availability of different payment options, the lack of accessibility to everyone "The Grammar Business as useful as it is, is however expensive and not available as trial." and lack of the availability of various features "Sometimes the features are not available automatically, and I have to upload the window I'm working on (even several times) to make it work."

Users also highlighted their concerns regarding the integration aspect of the AI writing assistants. Specifically, they complained about the limited integration of the tools with other software or platforms such as the absence of integration with Google Doc "Need google docs integration this would great for people who write online," desktop applications "It's very straightforward. Could be improved with more integrations like Microsoft PowerPoint or other desktop apps." and Microsoft products "integration with Outlook is not great and plagiarism check is not as expected."

For customer support attribute, responsiveness was the most commonly referenced category in the positive review as shown in Figure 6. The analysis of the associated keywords of this category, in Table 4, shows that "Communication" was the most cited, in the positive reviews. The interesting aspect of this feature is that users appreciated the capability of the tool in streamlining the internal communication among team members in the business "It makes communication between different groups so much easier. Especially when working with people who do not have English as a first language. This bridges the gap." As well as enhancing the external communication with customers through email writing, social media communication, and more efficient handling of customers' inquiries "We use it for improving the quality of our internal communication via emails, reports as well as external communication - newsletters and blog posts."

It is worth noting that in the customer support attribute, competence/assurance and empathy/relationship service categories received higher mention frequency compared to other categories in the negative reviews, as shown in Figure 6 and Table 4, indicating that there were significant concerns about the ability of the agent in resolving the issues at hand and creating positive customer experience "Customer support not very quick or helpful. Sometimes features does not work. It could definitely get better."

For the value for money attribute, Figure 6 reveals that value for money category was central to users' concern in the negative reviews. Although many users of different AI writing assistant appreciated the benefits and significant values of the tools, there was a recurring concern regarding the price of the software. It is also worth noting that when comparing the frequency of the "value for money"

related keywords, which reflect how the price of the software is perceived relative to its value in both positive and negative reviews, the frequency in the positive reviews is 147 as shown in Table 4. On the other hand, the frequency of keywords indicating users' complaints about the high price of AI writing assistant tools is 726 in the negative reviews. This finding reflects the user dissatisfaction with the current price model, suggesting a need for more flexible and affordable price plan.

Examples of users' feedback on the software price in the negative reviews: "My only gripe would be the monthly cost. It's a little pricey for the monthly rate, so I highly recommend choosing the yearly plan to save some money." "1. Upgrading account into premium is very pricey. 2. The offer service for free users are limited." Additionally, some users such as entrepreneurs or freelancers were opting for AI writing assistants' free version. This may suggest that the price is a major barrier to user satisfaction and adoption of the full version. For instance, "The free version should have more suggestions. Premium and business versions are a little bit costly. Other than these everything is AWESOME."

4.3. Text mining results

The third research question concerns identifying the attributes or features that are considered specific to AI writing assistant tools for business. The following sections classify these attributes into two categories: attributes that are common which represent general aspects that AI writing assistant tools users commented on and are applicable to various tools; and attributes that are specific to AI writing assistant tools.

4.3.1. Common attributes

The common dimensions that are critical for AI writing assistant tools users in evaluating their level of satisfaction include: price and payment plan (6 topics), customer service and support (4 topics), software update and expandability (1 topic), learning (1 topic), usefulness (2 topics), integration (3 topics), and accuracy (2 topics).

The analysis of the extracted topics from the negative reviews revealed that the most frequent attribute revolved around cost (C1), pricing (C2), payment plan (C3), and free and trial version (C4). Six of the generated topics in the dissatisfied reviews were primarily highlighting the price issue. This indicates that AI writing assistant tools users are highly concerned about the cost of the tool and affordability. Addressing this concern by providing more flexible payment options and improved trial version could enhance user satisfaction.

C1. "The only cons are the cost (my employer pays but we are a start up and every dollar counts) and other then that none so far."

C2. "pricing may be a barrier for some users, as it is a premium tool with a higher price point than some other content creation and management tools on the market."

C3. "I think the platform should at least reduce the price and adjust the payment plan"

C4. "I think the trial period is a bit short, especially for a small business to be able to completely use the product"

The second important attribute is customer service and support. Four of the extracted topics in the dissatisfied reviews emphasized this attribute. After analyzing users' reviews, we found that customer service and support includes activities such as answering customer inquiries (C5), solving their problems (C6), the availability of different contact options (Omnichannel support) (C7), time to response (C8), and support options for different customer categories (C9).

C5. "Customer service gave useless answers (clear cookies etc.)"

C6. " … .My latest issue is still yet to be resolved, which is the "web based product" … ."

C7. " … .The customer service department can not be reached by phone at all, you would only be able to reach them via email."

C8. "Perhaps customer service support can be improved as sometimes they are slow to respond."

C9. "The gap between the pro and free versions is too much. There is no education version available for students."

The third critical dimension is software update and expandability. The important aspects of the software update and expandability primarily include bug fixes (C10) and adding new features or improving existing one (C11 and C12).

C10. "Needs to update regularly some minor bugs need to be fixed."

C11. "I hope the makers are going to continue to improve it. Adding more languages to the list would be wonderful."

C12. "Some of its features and needed to be update."

The fourth prominent attribute is learning curve. The learning curve describes the learning performance of a task with gained experience (Martin et al., 2011). It also includes the time needed to complete a task and the likelihood of making mistakes. Users expect to become proficient with the tool in less time and effort. The steeper the curve, the harder it is to achieve proficiency (C13). It is obvious that learning curve matters as it influences how easily AI writing assistant tools users learn and adapt to the tool (C14).

C13. "There is a steeper learning curve because there are so many functions. This is not a complaint about the product. It takes some time to become proficient with all the tools."

C14. "the initial learning curve for some team members adapting to the tool was a minor hurdle. Once overcome, Grammarly Business proved highly beneficial."

Fifth, usefulness is one of the major concerns for AI writing assistant tools users. Usefulness describes the degree to which a software or a tool enhance the user job or performance (Davis, 1989). Usefulness reflects several aspects of a software such as ability to accomplish tasks more quickly, improves the work quality, makes the work easier, improve the job performance, and increase the productivity (Keil et al., 1995). In the present study, the users commented on different aspects of usefulness. Users, for example, emphasized the usefulness of the tool features that enable them to accomplish their tasks (C15), the usefulness of the generated content that improves the quality of the work (C16), and the missing features that, if available, could improve the productivity and make the job easier (C17).

C15. "nothing that I don't like about the software. It's pretty useful once you get through all the features and shortcuts to help in content writing."

C16. "Nothing so far; it's easy to use and delivers useful content."

C17. "It's only available for the English language. It would be very useful if it would be available in Spanish and Franch."

The sixth common attribute is integration. Three of the generated topics in the negative reviews reflected the integration issues. Users, for example, were primarily complaining about the tool's integration with different business apps or platforms (C18 and C19) and the integration process itself (C20).

C18. "The MS Office integration is not available for Apple users."

C19. "Seamless integration to various other platforms is still missing key."

C20. "Integration was pretty smooth but could be improved."

Finally, accuracy is one of the important attributes that received a significant attention from users. Two of the extracted topics from both the negative and positive reviews discussed this attribute. In the dissatisfied reviews, users highlighted the content accuracy issue (C21) and tool features accuracy such as suggestion, mistake detection, auto-correction, and spell check (C22 and C23).

C21. "Sometimes the accuracy of the generated content could be wrong and we need to fact check it again."

C22. "I wish the accuracy increase more on spelling checker and auto correction"

C23. "Sometimes the edits/suggestions that it makes can be wrong. But as with anything, you should always double check the accuracy"

In the positive reviews, on the other hand, users discussed the positive aspects of the accuracy such as the content accuracy (C24) and features accuracy (C25).

C24. "I loved all the features of Jarvis. and the best one I love recreation of the content and accuracy of content."

C25. "It's accuracy for grammar, punctuation and spelling mistakes. This really helps my documents be more professional."

Worth noting that the identified attributes using topic modeling algorithm closely align with those extracted through text analysis part in this paper. This convergence highlights the validity of the identified attributes and the strength of these factors in determining user satisfaction and dissatisfaction. Consequently, we categorized the common dimensions derived from the topic modeling under a broader attributes identified in the attributes dictionary, which was developed based on established theoretical frameworks (see Table 2). The synthesis of the derived attributes into broader attributes was guided by their alignment with the predefined attributes, including their associated categories and linked keywords. For instance, the topics related to the price and payment plan were categorized under the value for money attribute. Customer service and support related topics such as answering customer inquiries, solving their problems, the availability of different contact options, time to response, and support options for different customer categories were mapped into customer support attribute. Software update and expandability topic was classified under functionality attribute as this topic is aligned with system/software quality category of functionality attribute. Learning topic, on the other hand, was grouped under ease of use attribute. Further, usefulness and integration topics were aligned with functionality attribute, as these topics are associated with the items of system/software quality category of functionality attribute as shown in Table 2. Accuracy topic was also grouped under functionality attribute. This topic is linked to the performance items category of functionality attribute. This mapping process enhances interpretability of user feedback as well as the synthesis into conceptually meaningful themes.

4.3.2. AI writing assistant tools specific attributes

The specific attributes that are critical for AI writing assistant tools users to evaluate their level of satisfaction include: AI writing assistant tools specific functionalities such as editing capability (2 topics), content style (1 topic), English language (1 topic), suggestion and contextual suggestions feature (4 topics), SEO function (1 topic), checking functionality (1 topic), AI-generated content (1 topic), article tone (1 topic), Auto-Completion feature (1 topic), templates (1 topic), correction functionality (1 topic), ignore option (1 topic), social media content (2 topics), text quality (1 topic), plagiarism (1 topic), research use (1 topic), UX and AI Advancements (1 topic), and creativity (1 topic).

The first specific attribute is editing capability. AI writing assistant tools offer distinctive features and capabilities such as content editing including correcting punctuation mistakes, grammar and spelling check, and writing style adjustment. As the content editing feature was one of the primary determinants of AI writing assistant tools' users' satisfaction, the following complaints shows the users' frustration:

C26. "some edits aren't always accurate for every purpose behind my writing. It is critical you look at all edits/suggestions before blindly making them."

C27. "It doesn't have multiple paraphrasing features for editing."

C28. "Another cons is there is no offline editing mode."

C29. "There is a lot of lag when editing large sections of text."

The second factor is English language. One of the major concern is AI writing assistant tools is multilingual support. This feature enables the users to generate or edit content in multiple languages.

C30. "It lacks some useful features like multilingual support."

C31. "Only English is supported. The tool won't work correctly if you write in a different language."

Suggestion and contextual suggestions feature is the third AI writing assistant tools specific dimension. One of the powerful feature of these tools is suggestion including word or sentence suggestion, grammar error suggestion, and idea suggestion. This feature also includes generating content or suggestions that are specific to a context. This feature topic is extracted from both the negative and positive reviews. The following comments reflect user dissatisfaction with this feature (C32 and C33), while other comments (C34 and C35) indicate user satisfaction.

C32. "Some suggestions are not to the point irrelevant suggestions"

C33. "The correction proposal sometimes does not match the context of the statement."

C34. "The tool can check on a context-specific writing and can also suggest a possible "rephrase"."

C35. "Not only does it suggest better ways to phrase things, but it also helps me brainstorm ideas and come up with new topics to write about."

Social media content is also a factor that extracted from both positive and negative reviews. AI writing assistant tools help in generating social media content, posts, suggesting ideas, and headlines. Some users found the tool is effective in creating social media content (C36), while others were dissatisfied, highlighting the limitation of the tools in this feature (C37).

C36. "Whether you need blog posts, product descriptions, or social media content, the platform has you covered. The AI-generated content is not only high-quality but also highly customizable, allowing you to tweak and edit the output to suit your needs."

C37. "It would be nice if they had a social media post option."

The other specific factors that are extracted from the dissatisfied reviews include: SEO function (C38) to optimize content that can be ranked higher in the results of search engine by creating content that is keyword-optimized, readable and well-structured, plagiarism free, and search intent oriented.

Checking functionality (C39) is used for grammar and spelling check, plagiarism check, content accuracy checks, and style and formatting check.

AI-generated content (C40) is a primary task of AI writing assistant tools including various media such as images, video, audio, and text that can be used in social media writing, article and product description, and marketing material.

Article tone (C41) is one capability of AI writing assistant tools that reflects the style of writing such as causal, formal, and any other tone.

Auto-Completion feature (C42) is a feature that enables users to quickly complete the task by suggesting words or phrases. Templates (C43) is a functionality that allows users to create content according to a specific format such as email, ad copy, and blog.

Correction functionality (C44) relates to the checking functionality that ensures the text is free from grammar, punctuation, and spelling errors.

Ignore option (C45) is a feature that enables the user to dismiss the suggestions made by the tool.

In addition, text quality (C46) is one of the primary functionality of AI writing assistant tools essential for producing high quality content. Finally, plagiarism checker (47) feature allows users to make sure the content is original.

C38. "It will be great to include features tool such as generating SEO meta descriptions or SEO keywords to help us generate high converting content."

C39. "I am only conscious of fact-checking the information generated against reliable sources. While AI isn't likely to fact-check against scientific literature right now, I think it will be an excellent addition in the future."

C40. "What I don't like about the tool is the fact that it limits customization and the AI-generated content sounds too robotic which I think can highly affect the quality and originality of the content as they may lack some human touch and can potentially affect the overall quality of marketing performances."

C41. "Would like some more variations on tone, specifically within professional writing."

C42. "Sometimes the auto complete content isn't perfect and need re-writing for grammar fixing."

C43. "I wish they could add marketing email templates as well instead of only press emails."

C44. "In my experience, spelling correction has always been on point but that's a given, grammar errors are what the real concern is and if it fails to deliver the right grammatical error then it is not of much use."

C45. "There is also no "ignore all" option for when it suggests something is wrong and it is not."

C46. "Though its a good tool, I found that the quality of text produced is not always as high as with ChatGPT. For this reason, I usually combine the two for maximum effect."

C47. "The content is produced is also not what you'd expect for the price. There's a lot of ending and I found a lot of plagiarism so beware if you're using this for SEO."

The specific attributes that are generated from the satisfied reviews include: research use (C48) reflects the ability of the tool to be used in academic writing style, UX and AI Advancements (C49) is the use of AI writing assistant tool for creating user experience (UX) content. Creativity (C50) reflects the ability of the tool to bring new and more ideas, produce compelling content, and provide unique insights on the content.

C48. "Since my research work needs a lot of original writing inspired by previous work, the content rephrase tool in Writesonic helped me so much, its results are really accurate and I rarely need to change or correct some words."

C49. "The UI/UX is very well thought about. When a user researches a particular topic, the entire dashboard gives the writer a complete overview of the topic. The Headings (h1,h2), SERP Analysis, Questions asked, Reddit, PAA."

C50. "Jasper's AI technology uses natural language processing to generate human-like content quickly and efficiently. 2. Increased Creativity: The AI algorithm helps to bring out new ideas by analyzing existing content and providing suggestions for improvement."

These results underscore the relevance of these specific attributes in shaping user feedback. These AI writing assistants' specific aspects emerged as a key driver of user satisfaction or dissatisfaction and are highly relevant to the tools functionality attribute. The completeness of the tools functionality aspects enables the user to perform their different activities including input, processing, and output (Elias & Lubua, 2024). Therefore, if a user struggle to use the tool to perform their operations, the tool is considered as not suitable for completing a given task. In this study, users placed a significant weight on the tools specific functionality attributes. This highlights the importance of functionality as a critical component for meeting user requirements and determining the tools usefulness (DeLone & McLean, 2016).

Importantly, these specific functionality attributes are not captured in the dictionary derived from the literature (see Table 2), which was developed to reflect generalizable attributes that are common across technologies. Thus, the results suggest that these specific dimensions offer meaningful extension to the dictionary attribute through the lens of technology-specific functionality aspects to better capture the nuances of user expectations and experience in the context of AI writing assistant tools.

5. Discussion

5.1. Assessing user satisfaction with different AI writing assistant tools

The six AI writing assistant tools are analyzed and compared based on five aspects: overall rating, EOU, functionality, value for money, and customer support. The overall results of the Dunn post hoc test showed that Jasper and Writesonic outperformed than most of the tools in overall users' ratings and functionality attributes. Jasper also rated the highest in customer support attribute, while Grammarly and Writesonic outperformed in ease of use attribute.

The significant differences in the AI writing assistant tools users' ratings across the identified attributes suggest that users may experience distinct satisfaction levels in terms of EOU, functionality, value for money, and customer support depending on the AI writing assistant tool used. Consequently, influencing user satisfaction and tool selection.

In term of value for money, on the other hand, the result indicates that there is no enough evidence to conclude that most of the tools significantly differ in this attribute. This may indicate that the users perceive the AI writing assistant tools as providing similar value relative to their price.

5.2. Attributes that leading to user satisfaction and dissatisfaction

The results of text analysis (frequency of mention) show that the key determinants of the AI writing assistant tools user satisfaction are ease of use, functionality, customer support, and value for money. The findings reveal that these attributes are dual factors, meaning that they are leading to both user satisfaction and dissatisfaction, the result indicates that ease of use is the most attribute leading to user satisfaction. Obviously, without ease of use and user friendly tools, the unique advantages will not be materialized. The findings are in line with TAM which suggests that ease of use plays a vital role in technology adoption and satisfaction (Davis, 1989).

The results, on the other hand, show that value for money is the most leading factor of AI writing assistant tools user dissatisfaction. Although the users of different AI writing assistant value the benefits and features of the tools, there is recurring concern regarding the price of the software. Prior studies indicate that charging a reasonable price significantly contributes toward customer satisfaction and loyalty (e.g., Hanif et al., 2010). Thus, this suggests the importance of offering more fair price and flexible payment options.

Further, the results demonstrate that the source of satisfaction and dissatisfaction is different at the attributes level, categories level, and keywords level, which suggest that it is essential for the tools providers understand the causes leading to satisfaction and dissatisfaction at these different levels.

5.3. Investigating the key AI writing assistant tools related attributes

Listening to AI writing assistant tools users voice is the initial and critical step in the tools improvement planning. In turn, determining user satisfaction and adoption attributes and their roles in the satisfaction and dissatisfaction of the users offers a reference frame for the providers to evaluate their tools performance. Ideally, AI writing assistant tools providers should pay attention to the full spectrum of the identified attributes from both satisfied and dissatisfied reviews. These critical attributes embrace common factors, which represent general aspects that AI writing assistant tools and are applicable to various tools; and factors that are specific to AI writing assistants.

The general aspects of AI writing assistant tools are essential for enhancing user experience and satisfaction. For instance, software price is a major concern in the dissatisfied reviews. It has received a significant attention by users. Although they appreciate the value of the AI writing assistant tool and believe it is worth the price, they seek for more cost-effective tools and flexible payment plan. User's dissatisfaction with the tools price and payment plan has also been shown in text analysis section (Table 4) in this paper. This repeated instance of dissatisfaction highlight a consistent concern among AI writing assistant users. This means that the price has a major role in user satisfaction and adoption. Thus, it is advisable for the providers to further evaluate their solutions in terms of price and flexible payment options. Additionally, it is suggested that providers offer various strategies and business models

such free trial, freemium, and premium considering the supported software features with each business model to enhance customer retention and loyalty (Kim et al., 2018).

Further, customer service and support reflects the users' preference of offering multiple channels to contact AI writing assistant provider "The customer service department can not be reached by phone at all, you would only be able to reach them via email. When visiting the website there is no Customer service phone number or fax number only an email address." This comment reflects the need for solving problems promptly through the availability and promptness of communication channels. Multiple ways to communicate with the representative provide users flexibility, ease in reaching a support employee, and security (Yang & Fang, 2004).

Furthermore, update and expandability reflect the software technical integrity in terms of maintenance for or extend the software functionality. Previous related literature demonstrates the potential positive impact of software updates on users' satisfaction and use (e.g., Fleischmann et al., 2016). This factor presents the importance of understanding the user reactions over time. For providers, users with high satisfaction level are especially desirable goal as they are loyal and repeating customer can contribute to business long term profitability.

Learning attribute describes the progress in user's ability to acquire new skills through practice over time (Lallé et al., 2015). In the dissatisfied reviews, some comments show that there is a learning curve and users need adequate time to understand the tool and its different components before fully making use of its potential efficiently. Thus, the availability of proper training and sufficient tutorials in different languages and help resources on how to use the software can minimize the learning curve issue.

Usefulness, on the other hand, describes the extent to which the software enhances the individual's job performance (Davis, 1989). This indicates that AI writing assistant tools must include useful components and features that enable users to achieve their objectives. Integration attribute ensures the ability of different systems to communicate and flow data seamlessly. Users' comments, in this study, reveal the need for the AI writing assistants to integrate with various sources such other relevant tools, writing platforms, websites, google documents and tools, Microsoft products, spreadsheet, etc. to streamline the writing and editing process. The integration feature makes the AI writing assistants a valuable tool for businesses and it is a critical factor that contributes to user satisfaction (Bender, 2020).

Accuracy reflects the tools ability to generate accurate content. The accuracy aspects that are highlighted by AI writing assistant tools users include the accuracy of the generated content, free from different types of errors such as grammar, punctuation, and spelling errors, the accuracy of the suggestions and editing correction, sentence structure, and corrects with appropriate style and tones. Prior studies provide evidences about the significant effect of the accuracy on customer satisfaction in information system domain (e.g., Kim et al., 2021).

Our results, on the other hand, suggest that the users of AI writing assistant tools are concerned about 18 key specific attributes: AI writing assistant tools specific functionalities include editing capability, content style, English language, suggestion and contextual suggestions feature, SEO function, checking functionality, AI-generated content, article tone, auto-completion feature, templates, correction functionality, ignore option, social media content, text quality, plagiarism, research use, UX and AI Advancements, and creativity.

Editing capability, checking functionality, suggestion and contextual suggestions are prominent determinant for users' adoption and satisfaction. Well-developed AI writing assistant helps creating high-quality content, catching and fixing different types of errors such as grammar mistakes, spelling and contextual spelling errors, inconsistent terminology, and content style errors whether the writing is for social media writing, UX writing, or academic/professional writing. These attributes are highly associated with and affect content style, SEO function, article tone, auto-completion, template, correction functionality, ignore option, social media content, text quality, plagiarism, research use, and UX and AI Advancements. All these features ensure the writing is clear, accurate, creative, consistent, optimized for search engine, and customized according to the user specific need and writing style.

Additionally, AI-generated content, creativity, and language are also related distinctive and critical features in enhancing the usefulness of AI writing assistants. Users demand creative content that seems natural, more human, and less AI-generated. They also require multi-language feature and not just English.

Worth noting that most of these specific attributes are extracted from the dissatisfied reviews. However, this does not mean that users are dissatisfied in the positive reviews. AI writing assistants should pay closer attention to these attributes as they play a major role in user satisfaction and acceptance of the tools.

Additionally, the derived attributes through topic modeling are closely aligned with the identified attributes using frequency of mention method. This consistency emphasizes their validity and relevance in influencing user satisfaction and shaping their experience.

5.4. Implications

5.4.1. Theoretical implications

Theoretically, the literature review in this study reveals that no study was found specifically devoted to assess the pros and cons in using specific writing assistance tools. The empirical part of this study adds to the body of knowledge in AI writing assistant tools by utilizing quantitative multi-method research design to review these tools based on actual users' positive and negative reviews. Users reviews based on practical use can offer valuable information and insights into a product or service (Guo et al., 2017). Specifically, the quantitative multi-method research design in this study comprises the Kruskal-Wallis test to gain in depth numerical insights by exploring the significant differences among AI writing assistant tools based on attributes such as EOU, functionality, value for money, customer support, and overall rating. This analysis offers a deeper understanding of the strength and weaknesses aspects of AI writing assistant tools. The research design also includes the quantitative text analysis using frequency analysis of the identified attributes and topic modeling to extract latent attributes. This methodological approach provides a replicable framework for researchers for examining other enterprise software domains.

The current study also adds to the body of knowledge in business AI writing assistant tools domain by identifying several attributes that influence user satisfaction. This study also offers relatively a comprehensive pool of attributes and sub-attributes for evaluating AI writing assistant tools and can serve as a foundation for future research to further validate these attributes. Moreover, the current study identifies the primary determinants of AI writing assistant tools satisfaction and dissatisfaction as well as the common attributes and specific attributes that are applicable to this context. These derived attributes can also inform theoretical models such as technology acceptance and information system success models.

5.4.2. Practical implications

Practically, this study offers important insights into what drive users' satisfaction and dissatisfaction for different stakeholders. Developers of AI writing assistant tools, for example, can gain better understanding of tools key attributes and functionalities that influence user satisfaction and adoption. For instance, price dissatisfaction is a recurring theme in user reviews. This suggests that dissatisfaction may stem from the communication issue of the tool's delivered values. Ensuring that users understand the tools available features and recognize the benefits and outcomes of the tools can help mitigate this issue. Additionally, offering flexible price models such as modular pricing plan may reduce dissatisfaction and frustration from paying for features they do not need or use. Understanding users' expectations and needs will place the AI writing assistant tools providers in a favorable position to track and enhance their business performance.

The derived attributes in this study could also help AI writing assistant tools providers to identify the strengths and weaknesses of these tools. Worth noting that most of the specific attributes derived from topic modeling were extracted from the negative reviews. This shed the lights on the tools performance bottlenecks and software quality aspects that developers need to address to stay competitive. Although superior AI writing assistant tool may not necessarily delight customers, tools providers must still meet industry standards and address any weaknesses in these attributes to avoid user dissatisfaction.

The findings from this study also offers valuable implications for corporate IT decision-makers. Specifically, the synthesized themes from user reviews help IT decision makers identify the tools that match to organization's needs. Additionally, understanding what attributes that drive user dissatisfaction

experienced by peer organizations allows IT leaders to anticipate the potential issues during the tool selection process. For instance, the specific AI writing assistant tools functionality attributes that were extracted from user reviews serve as a guidance for selecting the tool that align with organizational priorities. Therefore, managers should consider a structured approach involving key stakeholders (e.g., the tools users, IT, and executives) when selecting among AI writing assistant tools. They must weigh various strategic and operational factors for more transparent and justifiable selection decision. For instance, ease of use encourages widespread adoption across departments, whereas functionality aligns the tool with the intended tasks and is crucial for overall tool success and user satisfaction but may come at the cost of customizability. Similarly, customer support plays a key role in reducing the downtime and productivity loss. These key considerations should be navigated using a weighted decision matrix approach that allows aligning the tools with the organizational strategic and operational factors.

5.5. Limitation and future work

This study has some limitations that provide avenues for future research opportunities. The online customer review sites provide an opportunity for businesses and researchers to understand customer user perception and satisfaction with various AI writing assistant tools. In this study, we gathered the data from Software advice website as it provides both quantitative and qualitative data for various AI writing assistant tools and we excluded any tools that had fewer than 50 reviews from further analysis to minimize the impact of incomplete data. Thus, the findings may not be fully representative to other AI writing assistant tools. Future research my benefit from the longitudinal study to monitor the user preferences, expectations, and satisfaction of AI writing assistant tools various attributes and categories identified in the current study.

Further, this study is a quantitative multi-method research design, which employs Kruskal-Wallis statistical analysis with quantitative text analysis and topic modeling. The findings of the Kruskal-Wallis test provide valuable insights to AI writing assistants tools providers. However, future studies can extend this study and consider analyzing the impact of other factors such as user switching reasons to other tools on user satisfaction.

Additionally, the topic labeling process was manual and the interpretability of the extracted topics may be influenced by human judgement. Therefore, we recommend further validation in future studies. In the topic modeling results, On the other hand, the identified attributes are not mutually exclusive. This means some attributes may be further grouped under broader attributes. For instance, checking functionality, correction functionality, and text quality can be grouped as one attribute: accuracy and reliability. Future studies can also utilize the questionnaire survey method to further validate the identified attributes.

6. Conclusion

This study analyzes six popular AI writing assistant tools for business to assess user satisfaction across various attributes including user overall ratings, functionality, value for money, EOU, and customer support. This paper provides a reference point to business users who require knowledge about these tools to select the most suitable one for their businesses. We also develop a dictionary based on well-established models in the area of information systems to identify the key attributes and related categories and keywords or items that influence user satisfaction and adoption. This dictionary provides useful insights to AI writing assistant tools providers to further understand the factors that drive user satisfaction and dissatisfaction and optimize their products, enhance user experience, and increase the satisfaction and adoption rate.

This study also uncovers various common and AI writing assistant tools specific attributes from the pros and cons users' comments using topic modeling algorithm. By leveraging these attributes, tools providers can make data driven decisions to further improve their products and competitive position.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author contributions

CRediT: **Yousra Harb**: Conceptualization, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing; **Sérgio Moro**: Conceptualization, Writing – original draft, Writing – review & editing; **Ayman Harb**: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The work by S. Moro was supported by the Fundação para a Ciência e Tecnologia (FCT) within projects UIDB/04466/2023 and UIDP/04466/2023.

ORCID

Yousra Harb http://orcid.org/0000-0002-0906-9165
Sérgio Moro http://orcid.org/0000-0002-4861-6686
Ayman Harb http://orcid.org/0000-0003-1725-5707

Data availability statement

Data from this research can be available from the corresponding authors upon request.

References

Alaofin, T. (2024). *A revolutionary artificial intelligence ChatGPT may soon take your jobs* (Paperback ed). Tunde Alaofin.

Aburass, S., & Abu Rumman, M. (2024). Authenticity in authorship: The Writer's Integrity framework for verifying human-generated text. *Ethics and Information Technology*, *26*(3), 62. https://doi.org/10.1007/s10676-024-09797-z

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management*, *37*(3), 99–110. https://doi.org/10.1016/j.ijinfomgt.2017.01.002

Al-Okaily, A., Teoh, A. P., & Al-Okaily, M. (2023). Evaluation of data analytics-oriented business intelligence technology effectiveness: An enterprise-level analysis. *Business Process Management Journal*, *29*(3), 777–800. https://doi.org/10.1108/BPMJ-10-2022-0546

Alexander, T. (2024). Proactive customer support: Re-architecting a customer support/relationship management software system leveraging predictive analysis/AI and machine learning. *Engineering: Open Access*, *2*(1), 39–50.

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*(November), 101473. https://doi.org/10.1016/j.tele.2020.101473

Anugrah, F. T. (2020). Effect of promotion and ease of use on customer satisfaction and loyalty on OVO application users. *Quantitative Economics and Management Studies*, *1*(1), 44–50. https://doi.org/10.35877/454RI.qems77

Barreto, F., Moharkar, L., Shirodkar, M., Sarode, V., Gonsalves, S., & Johns, A. (2023). February) Generative artificial intelligence: Opportunities and challenges of large language models. In *International conference on intelligent computing and networking* (pp. 545–553). Springer Nature Singapore.

Bender, B. (2020). The impact of integration on application success and customer satisfaction. *Business & Information Systems Engineering*, *62*(6), 515–533. https://doi.org/10.1007/s12599-020-00629-0

Blei, D. M., Ng, A. Y., & Edu, J. B. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Chowdhury, M., Hosseini, M. R., Martek, I., Edwards, D. J., & Wang, J. (2021). The effectiveness of web-based technology platforms in facilitating construction project collaboration: A qualitative analysis of 1,152 user reviews. *Journal of Information Technology in Construction*, *26*, 953–973. https://doi.org/10.36680/j.itcon.2021.051

Culotta, A., & Mattei, N. (2024). Use open source for safer generative AI experiments. *MIT Sloan Management Review*, *65*(2), 11–12.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Deep, S., & Zanke, P. (2024). Digital transformation strategy with CRM and AI for SMB's sustainable growth. *ESP Journal of Engineering & Technology Advancements (ESP-JETA)*, *4*(3), 9–22. https://doi.org/10.13140/RG.2.2.35510.51529

DeJeu, E. (2024). A Rhetorical approach for reimagining business writing instruction in the AI age. *Business and Professional Communication Quarterly*. https://doi.org/10.1177/23294906241255301

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

DeLone, W. H., & McLean, E. R. (2016). Information systems success measurement. *Foundations and Trends® in Information Systems*, *2*(1), 1–116. https://doi.org/10.1561/2900000005

Desai, P., Wang, H., Davis, L., Ullmann, T. M., & DiBrito, S. R. (2024). Bias perpetuates bias: ChatGPT learns gender inequities in academic surgery promotions. *Journal of Surgical Education*, *81*(11), 1553–1557. https://doi.org/10.1016/j.jsurg.2024.07.023

Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. *Computer Methods and Programs in Biomedicine*, *102*(1), 75–80. https://doi.org/10.1016/j.cmpb.2010.11.002

Elias, J. D., & Lubua, E. W. (2024). The impact of usability, functionality and reliability on users' satisfaction during library system adoption. *The Journal of Informatics*, *1*(1), 13–21.

Elias, N. F. (2012). Using a deductive approach in validating a measurement model. *Australian Journal of Basic and Applied Sciences*, *6*(8), 329–336.

Fleischmann, M., Amirpur, M., Grupp, T., Benlian, A., & Hess, T. (2016). The role of software updates in information systems continuance—An experimental study from a user perspective. *Decision Support Systems*, *83*(March), 83–96. https://doi.org/10.1016/j.dss.2015.12.010

George, A. S., & George, A. H. (2023). A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*, *1*(1), 9–23. https://doi.org/10.5281/zenodo.7644359

Giordano, V., Spada, I., Chiarello, F., & Fantoni, G. (2024). The impact of ChatGPT on human skills: A quantitative study on twitter data. *Technological Forecasting and Social Change*, *203*(June), 123389. https://doi.org/10.1016/j.techfore.2024.123389

Güngör, H. (2020). Creating value with artificial intelligence: A multi-stakeholder perspective. *Journal of Creating Value*, *6*(1), 72–85. https://doi.org/10.1177/2394964320921071

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent Dirichlet allocation. *Tourism Management*, *59*(April), 467–483. https://doi.org/10.1016/j.tourman.2016.09.013

Goh, F. (2024). *Innovate to elevate!: A journey through mindset and Gen AI to enhance customer experience* (Hardcover ed.). WSPC.

Garcia, M. B. (2025). ChatGPT as an Academic Writing Tool: Factors Influencing Researchers' Intention to Write Manuscripts Using Generative Artificial Intelligence. *International Journal of Human–Computer Interaction*, 1–15. https://doi.org/10.1080/10447318.2025.2499158

Hanif, M., Hafeez, S., & Riaz, A. (2010). Factors affecting customer satisfaction. *International Research Journal of Finance and Economics*, *60*(1), 44–52.

Harb, Y., & Shang, Y. (2022). A semi-auto text mining approach for literature review: An example from IT for entrepreneurship. *International Journal of Knowledge Management*,)*18*(1), 1–16. https://doi.org/10.4018/IJKM.291093

Harb, Y., Shang, Y., Odat, W., Zaitoun, A. (2024). A review of popular business intelligence tools. *MWAIS 2024 Proceedings*. 11. https://aisel.aisnet.org/mwais2024/11

Hayajneh, S., & Harb, Y. (2023). Understanding the continuous use of business intelligence: The case of Jordan. *Journal of Decision Systems*, 1–32. https://doi.org/10.1080/12460125.2023.2253587

Hong, S., Thong, J. Y., & Tam, K. Y. (2006). Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. *Decision Support Systems*, *42*(3), 1819–1834. https://doi.org/10.1016/j.dss.2006.03.009

Hosanagar, K., & Krishnan, R. (2024). Who profits the most from generative AI. ?*MIT Sloan Management Review*, *65*(3), 24–29.

Jayanti, R. K., & Ghosh, A. K. (1996). Service value determination: An integrative perspective. *Journal of Hospitality & Leisure Marketing*, *3*(4), 5–25. https://doi.org/10.1300/J150v03n04_02

Jo, H., & Bang, Y. (2023). Understanding continuance intention of enterprise resource planning (ERP): TOE, TAM, and IS success model. *Heliyon*, *9*(10), e21019. https://doi.org/10.1016/j.heliyon.2023.e21019

Kasirye, F. (2021). An overview of mixed and multi method research. *Advance Preprint*, https://doi.org/10.31124/advance.14681643.v1

Keil, M., Beranek, P. M., & Konsynski, B. R. (1995). Usefulness and ease of use: Field study evidence regarding task considerations. *Decision Support Systems*, *13*(1), 75–91. https://doi.org/10.1016/0167-9236(94)E0032-M

Kim, H. W., Chan, H. C., & Gupta, S. (2007). Value-based adoption of mobile internet: An empirical investigation. *Decision Support Systems*, *43*(1), 111–126. https://doi.org/10.1016/j.dss.2005.05.009

Kim, J., Choi, I., & Li, Q. (2021). Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches. *Sustainability*, *13*(11), 6165. https://doi.org/10.3390/su13116165

Kim, J., Lee, J., & Zo, H. (2018). Toward sustainable freemium software: The roles of user satisfaction and use context. *Journal of Electronic Commerce Research*, *19*(3), 201–222.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Lallé, S., Toker, D., Conati, C., & Carenini, G. (2015). March). Prediction of users' learning curves for adaptation while using an information visualization. *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 357–368).

Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Ledro, C., Nosella, A., & Dalla Pozza, I. (2023). Integration of AI in CRM: Challenges and guidelines. *Journal of Open Innovation: Technology, Market, and Complexity*, *9*(4), 100151. https://doi.org/10.1016/j.joitmc.2023.100151

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research*, *66*(4), 579–619. https://doi.org/10.2307/1170654

Lu, Y., Yang, Y., Zhao, Q., Zhang, C., & Li, T. J. J. (2024). AI assistance for UX: A Literature review through human-centered AI. *arXiv*. https://arxiv.org/abs/2402.06089

Narayanaswamy, A. (2024). Evolution of microsoft copilot. In *Microsoft Copilot for Windows 11: Understanding the AI-powered features in Windows 11* (pp. 1–24). Apress.

Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems*, *21*(4), 199–235. https://doi.org/10.1080/07421222.2005.11045823

Martin, B., Mitrovic, A., Koedinger, K. R., & Mathan, S. (2011). Evaluating and improving adaptive educational systems with learning curves. *User Modeling and User-Adapted Interaction*, *21*(3), 249–283. https://doi.org/10.1007/s11257-010-9084-2

Mustață, M.-A., Iliescu, D., Mavriș, E., Jude, C.-R., Bojor, L., Tudorache, P., Cîrdei, I.-A., Hrab, D.-E., Aluculesei, A.-C., Răpan, I., Dan-Șuteu, Ș-A., Roman, D., & Urseiu, C. T. (2025). ChatGPT-Assisted Decision-Making: An In-Depth Exploration of the Human–AI Interaction. *International Journal of Human–Computer Interaction*, 1–16. https://doi.org/10.1080/10447318.2025.2499165

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460–469. https://doi.org/10.1177/002224378001700405

Page, M. J., & Moher, D. (2017). Evaluations of the uptake and impact of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement and extensions: A scoping review. *Systematic Reviews*, *6*(1), 263. https://doi.org/10.1186/s13643-017-0663-8

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.

Park, J., Oh, C., & Kim, H. Y. (2024). AI vs. human-generated content and accounts on Instagram: User preferences, evaluations, and ethical considerations. *Technology in Society*, *79*(December), 102705. https://doi.org/10.1016/j.techsoc.2024.102705

Parra-Frutos, I. (2013). Testing homogeneity of variances with unequal sample sizes. *Computational Statistics*, *28*(3), 1269–1297. https://doi.org/10.1007/s00180-012-0353-x

Pasch, S., & Ha, S. Y. (2025). Human–AI Interaction and User Satisfaction: Empirical Evidence from Online Reviews of AI Products. *International Journal of Human–Computer Interaction*, 1–22. https://doi.org/10.1080/10447318.2025.2545463

Patel, K., Beeram, D., Ramamurthy, P., Garg, P., & Kumar, S. (2024). Ai-enhanced design: Revolutionizing methodologies and workflows. *Development (IJAIRD)*, *2*(1), 135–157.

Pattanayak, S. K. (2022). Generative AI for market analysis in business consulting: Revolutionizing data insights and competitive intelligence. *International Journal of Enhanced Research in Management & Computer Applications*, *11*(12), 74–86.

Peters, M. D., Wieder, B., Sutton, S. G., & Wakefield, J. (2016). Business intelligence systems use in performance measurement capabilities: Implications for enhanced competitive advantage. *International Journal of Accounting Information Systems*, *21*(June), 1–17. https://doi.org/10.1016/j.accinf.2016.03.001

Phillips, J., & Robie, C. (2024). Hacking the perfect score on high-stakes personality assessments with generative AI. *Personality and Individual Differences*, *231*(December), 112840. https://doi.org/10.1016/j.paid.2024.112840

Rafique, H., Almagrabi, A. O., Shamim, A., Anwar, F., & Bashir, A. K. (2020). Investigating the acceptance of mobile library applications with an extended technology acceptance model (TAM). *Computers & Education*, *145*(February), 103732. https://doi.org/10.1016/j.compedu.2019.103732

Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*, *12*, 26839–26874. https://doi.org/10.1109/ACCESS.2024.3365742

Rajaguru, R. (2016). Role of value for money and service quality on behavioural intention: A study of full service and low cost airlines. *Journal of Air Transport Management*, *53*(June), 114–122. https://doi.org/10.1016/j.jairtraman.2016.02.008

Retkowsky, J., Hafermalz, E., & Huysman, M. (2024). Managing a ChatGPT-empowered workforce: Understanding its affordances and side effects. *Business Horizons*, *67*(5), 511–523. https://doi.org/10.1016/j.bushor.2024.04.009

Roslan, F. A. B. M., & Ahmad, N. B. (2023). The rise of AI-powered voice assistants: Analyzing their transformative impact on modern customer service paradigms and consumer expectations. *Quarterly Journal of Emerging Technologies and Innovations*, *8*(3), 33–64.
Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society*, *75*(November), 102372. https://doi.org/10.1016/j.techsoc.2023.102372
Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591
Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, *61*(4), 66–83. https://doi.org/10.1177/0008125619862257
Silva, C. F., & Moro, S. (2021). Blockchain technology as an enabler of consumer trust: A text mining literature analysis. *Telematics and Informatics*, *60*(July), 101593. https://doi.org/10.1016/j.tele.2021.101593
Singh, H., & Miah, S. J. (2020). Smart education literature: A theoretical analysis. *Education and Information Technologies*, *25*(4), 3299–3328. https://doi.org/10.1007/s10639-020-10116-4
Tung, T. M. (2024). Opening up the workplace: The way AI tools are changing productivity. *Educational Administration: Theory and Practice*, *30*(3), 480–491. https://doi.org/10.53555/kuey.v30i3.1300
Tunmise, A., Amanda, K. (2024). Technological advances and business writing practices: Navigating the digital landscape. *EasyChair Preprints*, (13659). https://easychair.org/publications/preprint/HXVw/open
Usman, F. O., Eyo-Udo, N. L., Etukudoh, E. A., Odonkor, B., Ibeh, C. V., & Adegbola, A. (2024). A critical review of AI-driven strategies for entrepreneurial success. *International Journal of Management & Entrepreneurship Research*, *6*(1), 200–215. https://doi.org/10.51594/ijmer.v6i1.748
Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. *Proceedings of the 26th International Conference On Machine Learning*, ICML 2009 (pp. 1105–1112). Academic Press.
Wamba-Taguimdje, S. L., Wamba, S. F., Kamdjoug, J. R. K., & Wanko, C. E. T. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal*, *26*(7), 1893–1924. https://doi.org/10.1108/BPMJ-10-2019-0411
Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International Journal of Service Industry Management*, *15*(3), 302–326. https://doi.org/10.1108/09564230410540953
Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2000). *A conceptual framework for understanding e-service quality: Implications for future research and managerial practice*. (MSI Working Paper No. 00-115, pp. 1–49). Marketing Science Institute. Cambridge, MA.
Zhan, Z., Rahman, A. B. M. M., Cui, R (2023). Will ChatGPT remodify the education? The future of business education and research. In E. Y. Li., et al. (Eds.), *Proceedings of the International Conference on Electronic Business*., *23*, 773–778. ICEB'23 October 19–23, 2023.
Zhao, J., & Wang, X. (2024). Unleashing efficiency and insights: Exploring the potential applications and challenges of ChatGPT in accounting. *Journal of Corporate Accounting & Finance*, *35*(1), 269–276. https://doi.org/10.1002/jcaf.22663
Zumsteg, J. M., & Junn, C. (2023). Will ChatGPT match to your program? *American Journal of Physical Medicine & Rehabilitation*, *102*(6), 545–547. https://doi.org/10.1097/PHM.0000000000002238

About the authors

Yousra Harb is an Associate Professor of Information Systems in the Department of Information Technology at Yarmouk University, Jordan. She served as Assistant Dean for Quality Assurance at the Faculty of Information Technology and Computer Sciences. Her research interests include knowledge management, business analytics, and healthcare analytics.

Sérgio Moro is currently Dean of the School of Technology and Architecture at Iscte, Portugal. He is Full Professor of Information Systems, with focus on Decision Support Systems and Data Science. He has published more than 60 peer-reviewed articles in Q1 (Scopus/Web of Science) journals, with a Scopus h-index = 32.

Ayman Harb is an Associate Professor of Hospitality and Tourism Management at the University of Jordan. He served as Dean of the School of Tourism and Hospitality at the University of Jordan, Aqaba Branch. His work focuses on entrepreneurship and advancing sustainable tourism practices within the tourism and hospitality sectors.

Appendix

Table A1. Topics extracted from negative reviews (topic 1–36) and positive reviews (topic 37–42).

Topic number	Topic keywords	Topic label
1	user, editing, interface, slow, end, missing, low, basic, advanced, friendly	Limited Editing Capability
2	enough, style, example, save, case, place, description, mean, benefit, report	Content Style
3	find, app, pay, system, able, plan, access, unlimited, minor, pro	Payment Plan
4	used, google_doc, complain, count, quickly, frustrating, web, miss, phone, accuracy	Integration with Google Docs complain
5	know, information, show, cost, big, fully, apps, data, large, figure	Cost/Price
6	different, integration, made, got, provided, option, giving, else, happens, changed	Integration Options
7	come, future, fix, update, new, expected, manually, next, understandable, couple	Software Update and Expandability
8	language, english, suggested, kind, worked, wanted, mode, fixed, switch, window	English Language
9	suggest, usage, phrase, lower, checker, single, button, landing_page, recommendation, irrelevant	Suggestion Feature
10	accurate, pricey, catch, easily, read, drawback, afford, corrected, team, come_across	Accuracy
11	program, negative, day, name, seo, general, thought, post, purpose, trail	SEO Optimization
12	expensive, limited, premium, long, useful, subscription, monthly, price, number, function	Price
13	good, back, customer, account, pretty, service, looking, prefer, go, daily	Customer Service
14	student, company, grammatical, outline, easier, structure, getting, requires, include, almost	Student Support (Customer Support)
15	issue, perfect, quite, small, tell, checking, functionality, google, paste, major	Checking Functionality issue
16	product, difficult, edit, believe, microsoft, section, totally, completely, internet, voice	Integration with Microsoft Issue
17	problem, right, fact, generate, improve, result, generated, start, click, level	AI-Generated Content issue
18	price, high, website, pricing, point, helpful, especially, moment, given, provides	Price
19	give, change, context, provide, job, confusing, opinion, cheaper, personal, result	Contextual Suggestions
20	article, experience, ad, full, tone, longer, ability, project, choice, keywords	Article tone
21	want, bad, try, limit, written, auto, complete, paragraph, increase, learn	Auto-Completion Feature Issue
22	feel, added, satisfied, template, automatically, tried, creating, adding, long_form, two	Template
23	available, working, hard, part, trying, programs, belong, often, learning, convenient	Learning issue
24	service, application, output, go, run, necessary, feature, type, next_day, become	Service Feature
25	platform, incorrect, human, page, typing, poor, enjoy, add, ahead, unresponsive	Responsiveness Issues on the Platform
26	free, version, trial, wrong, paid, appreciate, advice, correctly, prior, belong	Trial Version
27	document, term, editor, turn, definitely, superior, context, skills, minor, able	Editor
28	suggestion, people, amazing, put, course, negative, update, go, generally, absolutely	Suggestion Feature
29	correction, take, understand, site, needed, advertised, response, avoid, adjusted, helpful	Correction Functionality
30	still, new, support, improved, currently, absolutely, response, knowledge, clearer, interaction	Support and Response
31	type, keep, choose, already, ignore, job, file, compared, remove, yet	Ignore Option
32	blog, amount, review, fast, happy, clear, downside, entire, need, suggest	Social Media
33	add, hope, real, meaning, create, direct, usefulness, common, should, summarize	Usefulness
34	credit, option, offer, worth, month, good, credits, approach, free_version, wish	Credit system
35	text, quality, idea, improvement, topic, share, community, feedback, based, type	Text Quality
36	complaint, specific, plagiarism, simply, test, longer, issue, developer, additions, ready	Plagiarism Checker
37	suggestion, text, document, free, professional, english, punctuation, user, used, improve	Suggestion Feature
38	article, blog, quality, idea, create, post, generate, template, user, topic	Social media content
39	research, brief, topic, cut, piece, minute, click, website, put, find	Research Use
40	ux, superb, chatgpt, chatsonic, attention, process, intelligence, stand, leap, bound	UX and AI Advancements
41	block, idea, utilize, exceptionally, creative, utilized, blunder, promise, report, extremely	Creativity
42	precision, beat, io, collect, downloading, supporting, optimizing, blessing, revolutionary, elsewise	Precision/Accuracy

Table A2. Dunn's post hoc test results.

Overall rating	Grammarly	Jasper	Copy.ai	Writesonic	Ginger	ProWritingAid
Grammarly	1	0.000002	0.696469	0.023537	0	0.00001
Jasper	0.000002	1	0.073456	1	0	0
Copy.ai	0.696469	0.073456	1	0.208075	0.010596	1
Writesonic	0.023537	1	0.208075	1	0	0
Ginger	0	0	0.010596	0	1	0
ProWritingAid	0.00001	0	1	0	0	1

EOU	Grammarly	Jasper	Copy.ai	Writesonic	Ginger	ProWritingAid
Grammarly	1	0.000015	1	0.407662	0	0
Jasper	0.000015	1	1	0	0	0
Copy.ai	1	1	1	1	0.017179	0.448137
Writesonic	0.407662	0	1	1	0	0
Ginger	0	0	0.017179	0	1	0.307521
ProWritingAid	0	0	0.448137	0	0.307521	1

Functionality	Grammarly	Jasper	Copy.ai	Writesonic	Ginger	ProWritingAid
Grammarly	1	0	0.013496	0	0	0.006504
Jasper	0	1	0.000138	1	0	0
Copy.ai	0.013496	0.000138	1	0.00013	1	0.341189
Writesonic	0	1	0.00013	1	0	0
Ginger	0	0	1	0	1	0.000005
ProWritingAid	0.006504	0	0.341189	0	0.000005	1

VOM	Grammarly	Jasper	Copy.ai	Writesonic	Ginger	ProWritingAid
Grammarly	1	0.269691	1	0.002003	0.000003	1
Jasper	0.269691	1	0.46956	1	0	1
Copy.ai	1	0.46956	1	0.274637	1	0.805751
Writesonic	0.002003	1	0.274637	1	0	1
Ginger	0.000003	0	1	0	1	0.000005
ProWritingAid	1	1	0.805751	1	0.000005	1

Customer support	Grammarly	Jasper	Copy.ai	Writesonic	Ginger	ProWritingAid
Grammarly	1	0	1	0	0	1
Jasper	0	1	0.975902	0.000006	0	0
Copy.ai	1	0.975902	1	1	0.035445	1
Writesonic	0	0.000006	1	1	0	0.000012
Ginger	0	0	0.035445	0	1	0.000073
ProWritingAid	1	0	1	0.000012	0.000073	1

Table A3. Kruskal-Wallis Test.

	Tools	Mean Rank
Overall Rating	Grammarly	5736.04
	Jasper	6074.67
	Copy.ai	4936.14
	Writesonic	5930.94
	Ginger	3310.47
	ProWritingAid	5186.96
EOU	Grammarly	5866.93
	Jasper	5563.65
	Copy.ai	5604.70
	Writesonic	5999.22
	Ginger	4088.72
	ProWritingAid	4730.22
Functionality	Grammarly	5649.66
	Jasper	6147.65
	Copy.ai	4199.69
	Writesonic	6151.68
	Ginger	3642.15
	ProWritingAid	5226.42
VOM	Grammarly	4428.25
	Jasper	4572.82
	Copy.ai	3672.39
	Writesonic	4658.75
	Ginger	3019.80
	ProWritingAid	4499.45
Customer Support	Grammarly	4015.08
	Jasper	4822.79
	Copy.ai	4026.00
	Writesonic	4446.17
	Ginger	2484.23
	ProWritingAid	3827.85