Towards Effective AI-Driven Reading Assistants: A Design Science Exploration

Conference Paper · June 2024

TOWARDS EFFECTIVE AI-DRIVEN READING ASSISTANTS: A DESIGN SCIENCE EXPLORATION

Completed Research Paper

Martin Hänel, University of Kassel, Kassel, Germany, haenel@uni-kassel.de

Thiemo Wambsganss, Bern University of Applied Sciences, Bern, Switzerland, thiemo.wambsganss@bfh.ch

Matthias Söllner, University of Kassel, Kassel, Germany, soellner@uni-kassel.de

Abstract

Recent advancements in AI have led to the introduction of tools that support researchers in scientific reading. Tools such as SciSpace have come to the forefront to assist users in reading scientific texts. However, there is an insufficient theoretical foundation on how to design these reading assistants as well as no evidence of their effects, especially given the recent progress. Specifically, past literature lacks insights on evaluated user requirements and design principles for the design of computer-assisted reading systems. Addressing these challenges, we draw on Design Science Research (DSR) to derive and evaluate a set of five design principles for computer-assisted reading systems. Building on flow theory as our theoretical lens, we develop and perform a first proof-of-concept evaluation of a prototypical implementation of our principles as a computer-assisted reading artifact. Our design principles support researchers and practitioners on how to design, evaluate, and compare their AI-reading tools more effectively.

Keywords: Reading Comprehension, Large Language Models, Reading Assistant, Design Science Research.

1 Introduction

The impact of artificial intelligence (AI) on human productivity is widely discussed in society and first studies underscore respective positive effects of AI (Wang et al., 2023). This enhancement is particularly notable among individuals engaged in cognitively demanding professions, such as white-collar or knowledge workers (Felten et al., 2023). Specifically, knowledge workers who leverage AI for assistance demonstrate markedly higher productivity levels. On average, they complete 12.2% more tasks, accomplish tasks 25.1% faster, and deliver results with a 40% increase in quality compared to their counterparts who do not utilize AI assistance (Dell'Acqua et al., 2023), suggesting that knowledge workers must adapt to collaborate effectively with AI systems.

An interesting rather disregarded problem domain is the support of humans in processing large amounts of complex information, e.g., when reading scientific papers or technical reports. The volume of information currently accessible worldwide, coupled with its rapid and continuous expansion (Sivarajah et al., 2017), needs to be constantly reanalyzed and understood by individuals. This situation highlights the problem that human ability and time to process such vast amounts of information are limited. Humans can only process 200–300 words per minute (Brysbaert, 2019). Whereas machines are able to process and handle large text data through advanced language models like GPT-4, which can deal with up to 128,000 tokens in its context window. This enables GPT models to analyze and interpret intricate patterns in text data that would be challenging for human cognition (Greco et al., 2023). Algorithms have the potential to change how we make sense of information (Davison et al., 2023). Nevertheless, to design such systems profound knowledge from both users and theory are required to ensure their long-term usage. Recognizing the importance of designing effective systems for reading support is crucial due to several key factors. First, as Heuer and Glassman (2023) note, creating accessible text tools represents a significant challenge at the intersection of Psychology, Human-Computer Interaction (HCI) and Artificial Intelligence (AI). The development of such tools is not just a complex technical endeavor bridging several disciplines; it also holds the potential to benefit many users, making the comprehension and navigation of text more accessible to a broader general audience or supporting visually impaired users. Second, the ever-increasing volume of scientific literature accentuates the need for new technology in this domain. Lo et al. (2023) underscores this point, highlighting that as the body of scientific work expands, so does the necessity for advanced tools to support the reading process. Although computer-assisted reading tools based on novel advances of AI bear high potential, past research still lacks an interdisciplinary design perspective. In fact, research bridge domains of AI, HCI and Psychology to investigate the design and effects of AI-based reading assistants is still scarce. Research lacks insights on user requirements paired with psychological literature to develop and design effective reading assistants. Information Systems is particularly suited to contribute to such research gaps and hence, we pose the research question (RQ):

RQ: What are design principles for AI-augmented reading systems that leverage advancements in artificial intelligence to enhance the reading flow and productivity of scientific readers?

To answer this question, we draw on the design science research (DSR) paradigm as outlined by (Hevner, 2007), to address the outlined challenges and investigate specific design knowledge for computer-assisted reading systems. DSR is particularly suited to contribute to multidisciplinary design challenges in a rigorous and transparent way (Gregor and Hevner, 2013). Our iterative process aims to generate and assess design knowledge, employing flow theory as our theoretical foundation (Nakamura and Csikszentmihalyi, 2009). To the best of our knowledge, there is no study that rigorously derives requirements from both scientific literature and potential users to derive design principles for AI-augmented reading systems based on AI.

Our investigation presents initial design principles and their evaluation, as per (Venable et al., 2016). The findings indicate that LLM-based computer-assisted reading systems offer substantial potential for augmenting researchers' scientific reading experience. These principles could significantly contribute to the advancement of computer-assisted reading systems design, particularly in integrating the latest capabilities of LLMs. They serve as a foundation informing computer-assisted reading system designers towards the newest design possibilities with Large Language Models. Our approach includes the derivation of four initial design principles, their implementation in a prototype, and a proof-of-concept evaluation to validate these principles.

We provide three contributions to the fields of computer-assisted reading system and design science:

1) We provide results from 20 user interviews on augmented reading artifacts based on flow theory.

2) We derive four distinct design principles for researchers and practitioners to effectively design, compare and evaluate computer-assisted reading system.

3) We provide a first proof of concept evaluation of the principles, shedding light on the potential but also the limitations of these systems for users. In this vein we also contribute flow theory in developing computer-assisted reading system and designs science research.

This paper is structured to cover several key areas. First, we explore the problems with scientific papers disrupting the reading flow and how computers can augment with reading scientific papers, focusing on flow theory behind our artifact. We explain why we chose this theory and its importance. Subsequently, we describe our method for generating design knowledge, in accordance with the three-cycle perspective introduced Hevner (2007).

2 Theoretical Background

2.1 Problems with scientific paper reading

The acquisition of knowledge is an essential task of knowledge worker (Kelloway and Barling, 2000). The time knowledge workers deal with this acquisition of knowledge in the form of reading impacts their performance and speed. Reading is an essential part of scientific work. Reading comprehension is the ability of a reader to interpret not just the explicit content of a text but also its underlying meanings, resulting in a mental representation (Ahmadi et al., 2013). This understanding is an intricate process, resulting from the interaction between the text and the reader. In addition, scientific literacy, which involves the ability to understand and critically evaluate scientific content, is essential for achieving one's goals in scientific fields (Britt et al., 2014). The process of reading and understanding scientific literature presents a multifaceted challenge for individuals engaged in scientific work. Howard et al. (2018) identifies a critical hurdle in this context: the development of the ability to effectively engage with, comprehend, and critique scholarly literature, a skill crucial for academic and professional development (Barr and Tagg, 1995). Cromley and Azevedo (2007) underscore this point, noting that individuals often find themselves lost amidst scientific jargon and complex methodologies, struggling to decipher the intrinsic value and applicability of the findings.

Scientific texts are often information-dense, with varying quality, and researchers are under significant time pressure. However, individuals need to constantly update and understand new knowledge, highlighting the importance of reading comprehension in this process (Kittur, 2017). Hindering the knowledge assimilation, defined as internalized information (Liew, 2007) through reading flow disruption problems leads to a productivity loss for scientists.

The problem of reading comprehension in scientific contexts is four-fold. Firstly, there is the challenge of understanding specialized terminologies, which are often field-specific and not part of general language use (Augustine and Greene, 2002). Today, many texts require interdisciplinary knowledge as fields like technology and marketing increasingly converge. Secondly, the lack of clarity about navigating the structure of scientific papers can confuse readers (de-la-Peña and Luque-Rojas, 2021). Thirdly, authors often express themselves too complexly, failing to adequately explain theoretical concepts. Fourthly, texts are often too lengthy, extending to 100 pages, which leads to cognitive overload, motivational issues, and problems with orientation. Individuals are experiencing cognitive overload due to the overwhelming amount of information they are exposed to, resulting in a decline in attention span and mental fatigue (Benselin and Ragsdell, 2016). These issues not only disrupt the reading flow but also impede the reader's ability to extract and interpret essential information effectively. The challenges often necessitate external searches and multiple readings, thereby breaking the reading flow and reducing efficiency. Addressing these challenges with computer-assisted reading systems is key to enhancing the comprehension and efficiency of reading scientific literature.

2.2 Computer-assisted reading

The research field of Computer-Assisted Reading was originally established by Atkinson and Hansen (1966). It involves a computer-controlled system for teaching reading of texts, tailored for individualized instruction. Atkinson's idea covers a range of reading tasks, initial vocabulary development, transfer to new vocabulary, and sentence comprehension, providing a holistic approach to reading education. In general, the research field dedicated to augmenting the reading of scientific papers is well-developed. For instance, Chen et al. (2011) delved into digital tools for augmenting students' reading comprehension. Similarly, Head et al. (2021) focused on the application of term annotations in scientific texts. Additionally, Lo et al. (2023) examined how information technology can assist scientists in their reading. There is a lot of work presenting such systems with different tasks e.g., summarization (Zhang et al., 2023), text personalization (Ashok et al., 2019), highlighting (Lee et al., 2016), to augment the reading process. OpenAI's recent addition of a document upload feature to GPT-4 is a good example of the growing importance of such systems that help with large textual data (OpenAI, 2023). This system function can be clustered into discovery, efficiency, comprehension, synthesis, and accessibility, which are critical for augmenting the reading perception of people with a system (Lo et al., 2023).

The research gap in the field emerges from the recent advancements in the field of LLMs, which offer significant opportunities to augment reading support. LLMs are increasingly being used in science (Schlagwein and Willcocks, 2023). This LLM-based reading systems have high potential in text comprehension (Bommasani et al., 2022). They can explain unfamiliar concepts (Kohnke et al., 2023). and excel at question-answering dialogues and in text summarization (Lee et al., 2022). Further, Porsdam Mann et al. (2023) have emphasized how text can be rewritten using LLMs to improve reading. We can design LLM-based reading systems which intervene between the reader and the text with the generation of flow optimized text or explanations which hinder reading disruptions, thereby enhancing the reading flow. While AI-reading systems use a variety of AI techniques, including machine learning and natural language processing (Forbus et al., 2007; Zhang et al., 2021), to tailor interventions to improve text comprehension (Di Mascio et al., 2010). LLM-augmented systems specifically leverage large language models to generate rewritten texts, which creates new opportunities for personalization (Porsdam Mann et al., 2023). Given these new personalization design capabilities of LLM-based reading systems, it is now an open question in the literature field to understand how we can design computer-assisted reading systems that effectively support the reading flow through LLM intervention. This involves determining how to calibrate such reading systems, in terms of comprehension, reading speed, and cognitive load (Kosch et al., 2020).

2.3 Flow theory as a kernel theory for the artifact design

We choose flow theory for the design of our artifact, because flow is relevant as it is the optimal condition you can reach in reading (Mcquillan and Conde, 1996). Flow is defined as a state of peak enjoyment, energetic focus, and creative concentration experienced (Nakamura and Csikszentmihalyi, 2009) derived from reading the text is an important factor in predicting its comprehension (Chevet et al., 2022). Furthermore, flow leads to desired work outcomes like task performance (Nadj et al., 2023). Unfortunately, this cognitive state can be interrupted by certain textual structures or phrasings, leading to neuronal processing conflicts and increased cognitive load (Glim et al., 2023). Reading difficult texts increased both voluntary and involuntary mind wandering (Soemer and Schiefele, 2019). However, some text characteristics also bring people into the flow state e.g., if the skill level of the text is appropriate to them (Towey, 2000). Through intervention we can create flow (Nakamura and Csikszentmihalyi, 2009). We argue that if we can adapt the text through flow-enhancing intervention towards individuals with computer-assisted reading systems, we can create this flow state based on the flow improving design principles leading to higher reading efficiency of scientific texts. Particularly we aim to design a user-centered system with design features that enable users to reach a flow state when reading and processing information. We do that, by employing flow theory as our theoretical lens for our design science methodology and how we collect, derive, and formulate requirements from users and literature and how we design and evaluate our system.

3 Research Methodology

Our study is guided by the Design Science Research (DSR) approach as proposed by Hevner (2007), emphasizing the creation of design knowledge through a systematic and structured process. Figure 1 shows the six consecutive steps that have been conducted to ensure design knowledge is created based on insights from the application domain and the knowledge base. We analyze requirements based on the environment (relevance) and knowledge base (rigor), derive four distinct design principles, and instantiate and evaluate them in a prototypical application. Overall, our project aims to contribute to a nascent design theory that gives explicit prescriptions for designing this class of artifacts (Gregor and Hevner, 2013). We followed a theory-driven design approach by grounding our research on flow theory (Nakamura and Csikszentmihalyi, 2009).

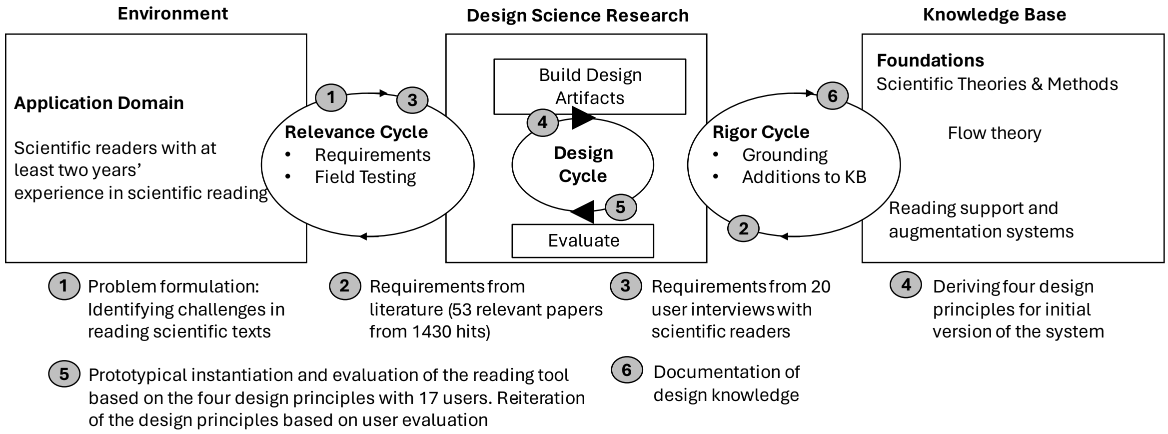

Figure 1. Overview of our design science research approach to rigorously derive and evaluated design knowledge for computer-assisted reading systems based on six consecutive steps adapted from Hevner (2007).

In the *first step* of our DSR project, we formulated the problem, identifying challenges in reading scientific texts within the realms of Information Systems, Psychology and Human-Computer Interaction research. This step was crucial in outlining the relevance of the practical problem, as detailed in the introduction and theoretical background of our work. In the *second step*, we conducted a systematic literature review following the procedure recommended by Vom Brocke et al. (2015). In our review, we found 53 relevant papers from a pool of 1430 hits. We selected Aisel, IEEE Xplore, ACM Digital Library, and arXiv as our chosen databases, focusing on the discipline of Information Systems, and conducted searches using the following keyword string: “Reading Instruction OR Computer-Assisted Instruction OR Reading Comprehension OR Reading Support OR Reading Intervention”. We excluded any articles that were either not relevant to computer-assisted reading systems or failed to meet established scientific quality standards. These papers provided us with a solid foundation for our problem domain. Based on these papers, we defined four overreaching literature issues that pose important requirements for designing computer-assisted reading tools. In the *third step*, we focused on better understanding the user by eliciting user stories (USs) through semi-structured interviews. We interviewed 20 users, aged between 23 to 41, enrolled in master’s or PhD programs, to learn and explore their experiences and challenges with scientific reading. The interviews, lasted on average 35 minutes and were based on a 15-item questionnaire. The questions in the interviews were informed by the prior literature we read in step two and were structured to cover user problems with reading scientific texts, experience with existing systems, and preferences regarding reading tool design. Based on the interviews, we derived eight representing user stories based on Cohn (2004). The *fourth step* involved the derivation of four distinct and explanatory design principles from the literature issues and user stories identified earlier. These principles were structured as per the guidelines suggested by Gregor et al. (2020), focusing on addressing the specific needs and challenges in scientific reading as revealed by our user research. In the *fifth step*, we developed a prototypical application of our reading support tool, embodying and instantiating our four design principles. We followed the evaluation framework proposed by Venable et al. (2016) and performed a formative and artificial evaluation of our initial version. Our goal in this proof-of-concept evaluation was to test the design principles instantiated in the prototype, learn about usability, and the overall effectiveness in enhancing the scientific reading experience. The feedback and insights gained from this evaluation were pivotal in refining the design and functionality of our tool, ensuring that it aligns closely with the user requirements and design principles we set out to address. Finally, in the *sixth step*, we revised the design principles based on the findings from step five and document our initial findings with this manuscript. With this paper, we presented the refined version of our design principles as well as the refined version of our prototypical application.

By following the DSR methodology, we aim to contribute a practical and user-centered tool that substantially improves the scientific reading process, informed by a rigorous approach, and grounded in real-world user needs and preferences.

4 Designing the Artifact

4.1 Deriving requirements from scientific literature and user interviews

To derive requirements from scientific literature, we initially focused our research on studies that demonstrate the successful implementation of augmented reading tools. Three broad areas for deriving requirements in literature were identified: educational technology, human-computer interaction, and reading literacy . We summarized similar topics of these contributions as literature issues (**LI**s) and formed four clusters.

We found that an important issue in literature on computer-assisted reading system encompassed the varying levels of knowledge (**LI1**) and reading proficiency of the user (Cao et al., 2015; Howard et al., 2018; Jáquez-Pérez and Villa-Maciel, 2021). Humans naturally come with different reading comprehension skills (Freed et al., 2017), but also with different objectives when reading and limitations (Alonzo, 2022). These factors should be regarded in the design of computer-assisted reading systems. Furthermore, it has been emphasized that for an optimal reading experience, an important literature issue is the influence of structure (**LI2**) towards reading (Akhondi et al., 2011; Kendeou and van den Broek, 2007; Paranjpe, 2009). For example, a well-structured text enhances reading comprehension.

Additionally, the use of systems for summarization (**LI3**) to expand the reading experience of the user (Ma et al., 2023; Zhang et al., 2023) was a discussed issue. Systems were used to summarize textual data in different ways. One example would be medical text summarization for a physician. Another important point in the literature was an interactive, iterative, and scaffolding dialog (**LI4**) (Feng et al., 2021; Gupta et al., 2020; Liang et al., 2023). This involves interacting with and extracting information from documents.

LI1	Varying levels of knowledge (Cao et al., 2015; Howard et al., 2018; Jáquez-Pérez and Villa-Maciel, 2021)
LI2	Structure of texts (Akhondi et al., 2011; Kendeou and van den Broek, 2007; Paranjpe, 2009)
LI3	Summarization (Ma et al., 2023; Zhang et al., 2023)
LI4	Interactive dialog (Feng et al., 2021; Gupta et al., 2020; Liang et al., 2023)

Table 1. Identified literature issues for computer-assisted reading systems.

Based on the derived **LI**s, we conducted twenty semi-structured interviews according to (Gläser and Laudel, 2010). The interviewees consisted of individuals who were potential users of computer-assisted reading systems in science. The interviewees were in mean = 27,95 years old (SD = 4,66). Of the 20 participants, 12 were female and 8 were male. The participants came from diverse scientific backgrounds, including linguistics, law, sustainability studies, psychology, and information systems. To gain impressions resulting from many years of scientific work, only participants with at least two years' experience in science were recruited for the interviews. The interviews underwent a qualitative content analysis process. Initially, they were coded to develop abstract categories. This coding utilized an open coding technique to ensure a standardized system for evaluation (Gläser and Laudel, 2010). We gathered 24 user stories (**US**) following (Cohn, 2004). The most frequently mentioned and significant user stories were selected in the next step. We aggregated the common user stories, which resulted in 8 final user stories displayed in Table 2.

US1	As a reader of scientific research articles, I would like a reading-assistance tool that provides concise summaries and key points for each paragraph, enabling me to easily grasp essential information and reduce my cognitive load.
US2	As a reader of scientific texts, I would like a reading-assistance tool that explains complicated content to me at a lower level of abstraction and with examples via a chat function, to make it easier to follow the text.
US3	As a reader of scientific texts, I would like a reading-assistance tool that transforms long sentences with difficult words into easier ones to improve my reading flow.
US4	As a reader of scientific texts, I would like a reading-assistance tool that highlights the important parts of a document for me, allowing for more efficient reading.
US5	As a reader of scientific texts, I would appreciate a reading-assistance tool that helps annotate unfamiliar words and content, enabling me to continue reading without interruption to search for definitions online.
US6	As a reader of scientific texts, I need a reading-assistance tool that translates words and sentences to another language, eliminating the need to use external translation services like DeepL.
US7	As a reader of scientific texts, I seek a reading-assistance tool to guide me to the parts of the text most relevant to my goals, saving time and keeping my focus on the important content.
US8	As a reader of scientific texts, I would like a reading-assistance tool that minimizes cognitive load to prevent reading interruptions and window switching. It should manage my motivation to keep me in the flow of reading.

Table 2. Aggregated user stories for computer-assisted reading systems.

4.2 Initial design principles

As illustrated, we have identified four **LI**s and **US**s as requirements for a computer-assisted reading system. The design principles and their grounding in the user stories and literature issues are displayed in Table 3. Our design principles were formulated in alignment with the flow theory (Gregor et al., 2020; Nakamura and Csikszentmihalyi, 2009). We argue that a reading system that instantiates our design principles will increase the reading flow of the user with the text. Based on these findings, we derived 4 preliminary design principles for the design of the system that makes scientific reading more efficient.

Principle of Simplicity: The **DP1** relates to system interventions that simplify text. Our interviews revealed that users lose their flow when encountering technical terms from an unfamiliar discipline and words that are not commonly used, as highlighted in **US3** and **US5**. Additionally, literature emphasizes in **LI1** that diverse knowledge among people, possibly including different languages mentioned in US6, can interrupt the reading flow. Therefore, **DP1** suggests that the system should offer translations and definitions to maintain a smooth reading experience.

Principle of Structure: In the user interviews, participants expressed difficulty in discerning the structure of texts **LI2**, noting that their reading flow is disrupted when they have to think about the purpose of a particular paragraph. Some users resort to creating their own highlighting **US4**. They also mentioned losing interest and flow when encountering content, they've already read. Additionally, users reported evaluating the importance of the text **US7**. This feedback indicates a need for a design principle that can both enhance text structure annotation and prioritize content for more efficient reading **DP2**.

Principle of Essential **DP3**: People often feel overwhelmed by the long sentences in a paper, which makes it hard for them to keep the flow and motivation , particularly when they lack the capacity to read the entire content, as indicated by **US3**, suggesting that their flow could be supported if the system helped in organizing overly long paragraphs and provide the essential content from the text in the form of short summaries and bullet points **US1**. The literature base on computer-assisted reading also supports this system design **LI3**.

Principle of Low Abstraction **DP4**: In the interviews, participants often mentioned being inspired by the latest technological developments. They expressed a desire to interact with the document **US8**, asking it if it contains any information relevant to their research, a feature inherent in **LI4**. Additionally, they were interested in receiving further explanation with examples on a low abstraction level for new difficult content from the system **US2**.

LI 1 - Varying levels of knowledge (Cao et al., 2015; Howard et al., 2018; Jáquez-Pérez and Villa-Maciel, 2021)	**DP 1** *Principle of Simplicity:* To enable designers and researchers to develop computer-assisted reading systems that facilitate the reading of scientific texts, they should leverage LLMs interventions for translation and definition of words and sentences that prioritize simplicity to enhance the reading flow.
US3 - As a reader of scientific texts, I would like a reading-assistance tool that transforms long sentences with difficult words into easier ones to improve my reading flow.	
US5 - As a reader of scientific texts, I would appreciate a reading-assistance tool that helps annotate unfamiliar words and content, enabling me to continue reading without interruption to search for definitions online.	
US6 - As a reader of scientific texts, I need a reading-assistance tool that translates words and sentences to another language, eliminating the need to use external translation services like DeepL.	
LI2 - Structure of texts (Akhondi et al., 2011; Kendeou and van den Broek, 2007; Paranjpe, 2009)	**DP 2** *Principle of Structure:* To enable designers and researchers to develop computer-assisted reading systems that facilitate the reading of scientific texts, they should be designed to structure the content effectively, emphasize key sections, and guide users through the document with LLM support, thereby efficiently promote the flow state of the reader.
US4 - As a reader of scientific texts, I would like a reading-assistance tool that highlights the important parts of a document for me, allowing for more efficient reading.	
US7 - As a reader of scientific texts, I seek a reading-assistance tool to guide me to the parts of the text most relevant to my goals, saving time and keeping my focus on the important content.	
LI3 - Summarization (Ma et al., 2023; Zhang et al., 2023) US1 - As a reader of scientific research articles, I would like a reading-assistance tool that provides concise summaries and bullet points for each paragraph, enabling me to easily grasp essential information and reduce my cognitive load. US3 - As a reader of scientific texts, I would like a reading-assistance tool that transforms long sentences with difficult words into easier ones to improve my reading flow.	**DP 3** *Principle of Essential:* To enable designers and researchers to develop computer-assisted reading systems that facilitate the reading of scientific texts, they should provide summaries and key points for each paragraph and shorten words and sentences with LLM intervention to retain the flow while reading.
LI4 - Interactive dialog (Feng et al., 2021; Gupta et al., 2020; Liang et al., 2023) US2 - As a reader of scientific texts, I would like a reading-assistance tool that explains complicated content to me at a lower level of abstraction and with examples via a chat function, to make it easier to follow the text. US8 - As a reader of scientific texts, I would like a reading-assistance tool that minimizes cognitive load to prevent reading interruptions and window switching. It should manage my motivation to keep me in the flow of reading.	**DP 4** *Principle of Low Abstraction:* To enable designers and researchers to develop computer-assisted reading systems that facilitate the reading of scientific texts, they should provide explanations at a low level of abstraction via LLMs to maintain the reader's flow.

Table 3. Initial design principles for a computer-assisted reading system.

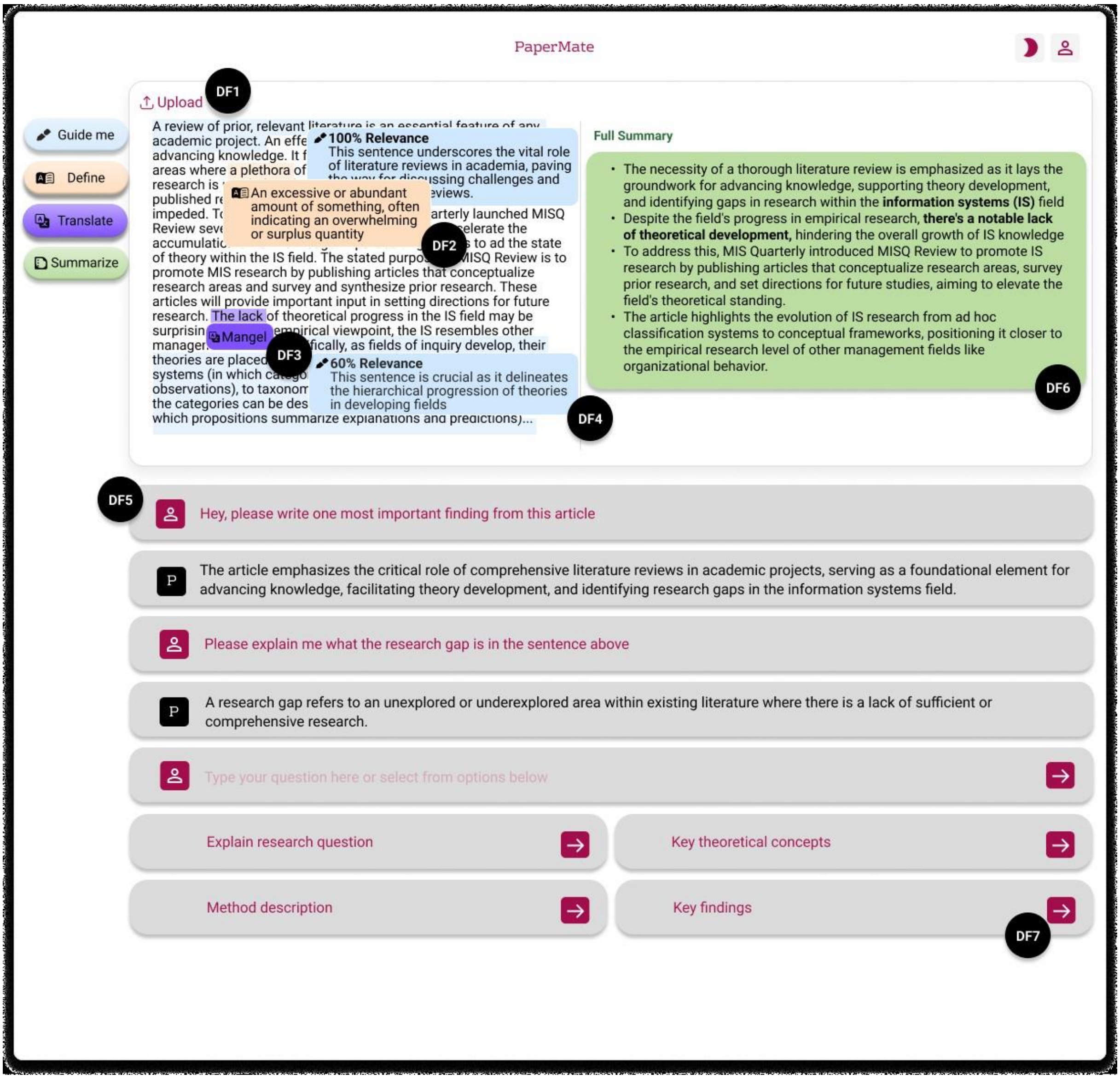

Figure 2. Initial prototype of the AI-reading assistant.

This paper introduces a prototype in Figure 2. Central to this prototype are seven key design features (DF) connected with the appropriate design principle (/**DP**), each contributing significantly to the handling, interpretation, and interaction with text-based information.

Ease of Document Import (**DF1/Basic Condition for all DPs**): Enhances user experience by simplifying the process of uploading and integrating PDF documents into the system for analysis. Clarification of Complex Terms (**DF2/DP1**): Aids users in understanding specialized jargon or intricate concepts by providing clear annotations and explanations. Accessibility of Multilingual Content (**DF3/DP1**): Improves user comprehension by translating foreign words, making texts accessible to non-native speakers or readers unfamiliar with certain languages. Efficient Information Retrieval (**DF4/DP2**): Helps users quickly locate and understand the significance of text sections, enhancing the efficiency of information extraction. User-Oriented Text Interaction (**DF5/DP2**): Allows users to interact with the text in a targeted manner, facilitating the location and comprehension of information relevant to their specific inquiries. Customizable Summarization (**DF6/DP3/DP5**): Offers users the flexibility to tailor summaries to their specific needs, enhancing the usability of the summarization function. Quick Access to Information (**DF7/DP2**): Provides users with immediate access to key information through predefined prompts, streamlining the information retrieval process.

4.3 Initial proof of concept evaluation

The aim of the evaluation was to confirm or refute our principles, developed based on scientific literature and user interviews, using both qualitative and quantitative methods. The findings could assist in modifying the design principles) or incorporating new ideas for potential design principles based on feedback from the participants. In conducting the evaluation, we adhere to the framework established by (Venable et al., 2016). The evaluation is performed as an ex-ante assessment using a prototype. We implemented an interactive prototype to assess the design, without a fully working backend. We utilized Marvel App, a tool for creating click-through prototypes, to simulate the frontend experience of the reading system.

We developed a prototype designed based on the identified principles. The participants interacted with the prototype for approximately 10 minutes. Afterwards, we showed them a questionnaire with our preliminary Design principles. We evaluated the principles with a 5-point Likert scale in the dimension of importance, usefulness, and robustness. Participants rated a statement on the scale, with '1' indicating strong disagreement (i.e., they did not believe the design principle is useful), '3' reflecting a neutral stance (i.e., they were neither in agreement nor disagreement), and '5' denoting strong agreement (i.e., they strongly believed that the design principle is useful). We provided an explanation in the questionnaire to ensure the knowledge of the respondents about design principles, to provide informed responses.

Giving the participants the chance to evaluate their choice through a text field for qualitative feedback. Furthermore we evaluated the artifact to demonstrate its utility and validity to solve the problem and be able to evaluate the system in terms of reading support (Gregor and Hevner, 2013). We measure the impact of the Design principles on the following constructs of the system ”Flow” (Thissen et al., 2018), “Fluency” (Lee et al., 2022),“Cognitive Load” (Gerven, 2003), ”Usefulness” and “Ease” (Venkatesh et al., 2003). The questions were adapted according to our reading system. At the end of the questionnaire three qualitative questions about the artifact were used. This qualitative feedback helped us to better understand how the users are perceiving the artifact. This qualitative feedback aimed to gain a deeper insight into users' perception of the artifact.

5 Results

5.1 Quantitative results

Seventeen potential users of the system evaluated the click-through prototype. All had experience in scientific reading for at least two years. The average age of the group was 30.19 years, with a standard deviation of 4.59. The participant group comprised ten males and seven females. They were recruited via mailing lists for participation in our evaluation. The results were analyzed both quantitatively, determining the mean and SD for each design principle and evaluation aspect, and qualitatively. The quantitative data from the questionnaire can be found in Table 4, while the qualitative findings are discussed in the text.

The evaluation demonstrated that experts generally view all design principles positively in terms of relevance, robustness, and usefulness. The mean values for these design principles are promising, especially when compared to the scale's midpoints. All design principles show strong significance compared to the neutral value of 3, with three design principles having absolute values greater than 4, highlighting their high appropriateness. **DP4**, while still above the neutral value, did not reach the score of 4, indicating it sparked considerable discussion among users.

Design Principle	Importance	Usefulness	Robustness	Overall Score
	The design principle is for designers who want to develop a user-centred reading system important	The design principle is for designers who want to develop a user-centred reading system useful	The design principle is for designers who want to develop a user-centered reading system robust	
1 Translation and definition of words and sentences	4.5 (SD=.816)	4.63 (SD=.806)	3.88 (SD=1.02)	4.33 (SD=.789)
2 Structure the content effectively, emphasize key sections, and guide users through the document	4.38 (SD=.806)	4.5 (SD=.816)	3.56 (SD=.963)	4.14 (SD=.750)
3 Provide summaries and key points for each paragraph, and shorten words and sentences	4.31 (SD=.870)	4.75 (SD=.680)	3.87 (SD=.885)	4.33 (SD=.677)
4 Dialog at a low level of abstraction	4.12 (SD=.957)	4.69 (SD=.478)	3.38 (SD=.873)	3.98 (SD=.818)

Table 4. Design principles evaluation results.

Scale: Strongly disagree=1, Disagree=2, Neither agree nor disagree=3, Agree=4, Strongly agree = 5

5.2 Qualitative results

DP1 was evaluated by the users as useful, with a mean of 4.19 (SD = 0.750), because they appreciated not having to switch tools for translation or definition while reading. One user commented, “I think it is very useful, since those are the operations, you usually have to use external tools for and switch applications.” However, they were also critical about the principle, mentioning, “Translations are rather a double-edged sword, helpful in the short run but not in the long run. I would not use them and would not recommend them to anyone who can understand the text without translation.” Comments like these highlight the need for all functions of such a system to be switchable off. There is no one-size-fits-all reading support system because people have different experiences, competencies, and goals while reading. **DP2** was evaluated by the users as very useful, with a mean of 4.44 (SD = 0.730). They particularly admired the potential of the principle to reduce their cognitive load, which received a mean score of 4.12 (SD = 1.02) while reading. However, despite the structure and guidance provided by such a system being valued for reducing cognitive load, there is an issue of trust among users. This is notable in the qualitative feedback, with comments like, “I kinda don't trust the guidance.” **DP3** performed very well in the dimensions of cognitive load (mean = 4.31; SD = 0.792), flow (mean = 4.06; SD = 0.928), and usefulness (mean = 4.43; SD = 0.629), with scores over 4. The focus on essential text content augmented people well, but they expressed that this design is only appreciated if they could see the initial, original text. They commented, “It would be good to be able to see the initial text in parallel. I don't trust the AI; I can compare directly.” **DP4** was discussed intensely during the user interviews. With a mean of 4.50 (SD = 0.632), the dialog design was evaluated as the most useful principle. There are individuals who admire this principle, and others who find it unnecessary for a computer-assisted reading system because it is not aligned with the reading flow (mean = 2.81; SD = 1.04). They argued, “The other design principles could already enhance text comprehension significantly.” Some people argued in favor, with comments like, “Absolutely necessary: Students use ChatPDF, ChatGPT; they expect these interactive features,” or “The interaction aspect of it really contributes to the ease of use of the system.” However, they also warned, “It is important that each answer refers to a point in the text that can be compared if necessary. The system should not re-hallucinate answers that cannot be deduced from the document as such.” We summarize this polarized discussion with a user comment: “Amazing function to get out of the text what I want out of it. Can't trust that it's robust.” Because of the ongoing mentions of trust, robustness, and explainability by the users towards the system and some evidence in the literature base (Kizilcec, 2016), we decide according to step six in Hevner (2007) to create and add: **DP5** Principle of Transparency and Explainability: To enable designers and researchers to develop computer-assisted reading systems that facilitate the reading of scientific texts, they should provide users automatically with the reasoning and changes the LLM-based system does with the text after reading to increase the readers trust in the system output and prevent active compare behavior of the reader while reading which destroys the reading flow.

To sum up, the value of such a system depends on its quality and transparency. A surprising finding is that even if the system were of high quality, people might not use it unless they can see the original text and the change history during reading. When these conditions are met, people have a positive attitude towards it, saying, “It could increase the efficiency of text analysis,” and “It’s a necessary tool for future science and business work to handle the increasing quantity and quality of information in every workplace.” If the system is well-designed, then a user stated, “I can imagine it being a more engaging process than just reading, and I also think that it can help focus your reading process on one application.”

6 Discussion And Conclusion

In this DSR project, we followed Hevner, 2007 to derive design knowledge in the form of design principles for computer-assisted reading systems. We rigorously derived requirements from 53 scientific papers (rigor) and from 20 semi-structured user interviews (relevance) to formulate a concise set of design principles for intelligent computer-assisted reading systems based on novel AI advances. We refined and evaluated those design principles by applying and instantiated them in an instantiated artifact with 17 users. We found that users using a reading support system based on our principles perceive them as important and useful. During our process, we evaluated the results of our design iteration and successfully finished our design process after the six steps.

Besides the prototypical artifact as a situated implementation of a computer-assisted reading system, we contribute design knowledge to the scientific knowledge base. We systematically deduced design knowledge as documented in our last step of the design research (Figure 1 in step 6). Due to the systematic procedure, we aimed at generating a satisfying design contribution (Gregory and Muntermann, 2014). The resulting design knowledge is not only valid for our specific case but can also be transferred to further use cases for computer-assisted reading system, e.g., to science, finance or technical reports. Due to this transferability of our design knowledge, our research does not only provide a Level 1 DSR contribution by showing a situated artifact implementation but also provides a preliminary form of a nascent design theory (Level 2 contribution) (Gregor and Hevner, 2013).

Past research that intersects the fields of Artificial Intelligence (AI), Human-Computer Interaction (HCI), and Psychology to explore the design and impacts of AI-based reading assistants is rather limited. There is a gap in research regarding the integration of user requirements with psychological insights for the development and design of effective reading assistants. This is especially evident in the field of Information Systems. With our work we have fill the gap. With this work we contribute to flow theory in system augmented reading (Chevet et al., 2022; Nakamura and Csikszentmihalyi, 2009). With our five design principles we contribute to the design science community following Gregor et al. (2020). This helps researcher and practitioners to effectively design, compare and evaluate computer-assisted reading system. In the long-term view this was the first step towards to contribute to a nascent design theory (Gregor and Hevner, 2013). We will continue working on this project aiming to build a programmed, naturalistic prototype to better evaluate a proof-of-value, comparative evaluation in laboratory settings according to Venable et al. (2016).

However, our research also faces several limitations. For the aim of this study, we focused our research on users recruited from our university and environment. Our evaluation was conducted with mostly students only, which usually have a certain experience in reading complex texts such as scientific papers compared to undergraduates. Even though it is reasonable to assume that the transferability to other cases is possible without major changes, we cannot prove it with our research design. Regarding the first implementation of our design principles in the prototypical application, we only implemented a prototype without a fully working backend. Nevertheless, we believe that a fully functional system might provide an even better usage for users. Also, our evaluate is limited to a first proof-of-concept evaluation with a low number of users. Future research is needed to evaluate the design of computer-assisted reading system with larger numbers of users over a longer period of time on behavioral outcomes.

This research illuminates the path for future development of AI-based reading assistants by providing empirically tested design principles and highlighting the importance of user-centered design considerations. By addressing the lack of theoretical and practical guidelines in the field, this work paves the way for more effective, transparent, and user-trusted AI reading tools, crucial for managing the ever-increasing volume and complexity of information across various domains.

References

Ahmadi, M.R., Ismail, H.N. and Abdullah, M.K.K. (2013), “The Importance of Metacognitive Reading Strategy Awareness in Reading Comprehension”, *English Language Teaching*, Canadian Center of Science and Education, Vol. 6 No. 10, pp. 235–244.

Akhondi, M., Malayeri, F.A. and Samad, A.A. (2011), “How to Teach Expository Text Structure to Facilitate Reading Comprehension”, *The Reading Teacher*, Vol. 64 No. 5, pp. 368–372, doi: 10.1598/RT.64.5.9.

Alonzo, O. (2022), “The use of automatic text simplification to provide reading assistance to deaf and hard-of-hearing individuals in computing fields”, *ACM SIGACCESS Accessibility and Computing*, No. 132, p. 3:1, doi: 10.1145/3523265.3523268.

Ashok, V., Billah, S.M., Borodin, Y. and Ramakrishnan, I. (2019), “Auto-Suggesting Browsing Actions for Personalized Web Screen Reading”, *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, Association for Computing Machinery, New York, NY, USA, pp. 252–260, doi: 10.1145/3320435.3320460.

Atkinson, R.C. and Hansen, D.N. (1966), “Computer-Assisted Instruction in Initial Reading: The Stanford Project”, *Reading Research Quarterly*, [Wiley, International Reading Association], Vol. 2 No. 1, pp. 5–25, doi: 10.2307/747037.

Augustine, S. and Greene, C. (2002), “Discovering how students search a library web site: a usability case study”, *College & Research Libraries*, Vol. 63 No. 4, pp. 354–365.

Barr, R.B. and Tagg, J. (1995), “From Teaching to Learning — A New Paradigm For Undergraduate Education”, *Change: The Magazine of Higher Learning*, Routledge, Vol. 27 No. 6, pp. 12–26, doi: 10.1080/00091383.1995.10544672.

Benselin, J.C. and Ragsdell, G. (2016), “Information overload: The differences that age makes”, *Journal of Librarianship and Information Science*, SAGE Publications Ltd, Vol. 48 No. 3, pp. 284–297, doi: 10.1177/0961000614566341.

Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., *et al.* (2022), “On the Opportunities and Risks of Foundation Models”, arXiv, 12 July.

Britt, M.A., Richter, T. and Rouet, J.-F. (2014), “Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information”, *Educational Psychologist*, Routledge, Vol. 49 No. 2, pp. 104–122, doi: 10.1080/00461520.2014.916217.

Brysbaert, M. (2019), “How many words do we read per minute? A review and meta-analysis of reading rate”, *Journal of Memory and Language*, Vol. 109, p. 104047, doi: 10.1016/j.jml.2019.104047.

Cao, M., Tian, J., Cheng, D., Liu, J. and Sun, X. (2015), “What Makes it Difficult to Understand a Scientific Literature?”, arXiv, 4 December, doi: 10.48550/arXiv.1512.01409.

Chen, N.-S., Teng, D.C.-E., Lee, C.-H., and Kinshuk. (2011), “Augmenting paper-based reading activity with direct access to digital materials and scaffolded questioning”, *Computers & Education,* Vol. 57 No. 2, pp. 1705–1715, doi: 10.1016/j.compedu.2011.03.013.

Chevet, G., Baccino, T., Vinter, A. and Drai-Zerbib, V. (2022), “What breaks the flow of reading? A study on characteristics of attentional disruption during digital reading”, *Frontiers in Psychology,* Vol. 13.

Cohn, M. (2004), *User Stories Applied: For Agile Software Development,* Addison-Wesley Professional.

Cromley, J.G. and Azevedo, R. (2007), “Testing and refining the direct and inferential mediation model of reading comprehension”, *Journal of Educational Psychology,* American Psychological Association, US, Vol. 99, pp. 311–325, doi: 10.1037/0022-0663.99.2.311.

Davison, R.M., Laumer, S., Tarafdar, M. and Wong, L.H.M. (2023), “Pickled eggs: Generative AI as research assistant or co-author?”, *Information Systems Journal,* Vol. 33 No. 5, pp. 989–994, doi: 10.1111/isj.12455.

Dell’Acqua, F., McFowland, E., Mollick, E.R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., *et al.* (2023), “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality”, SSRN Scholarly Paper, Rochester, NY, 15 September, doi: 10.2139/ssrn.4573321.

Di Mascio, T., Gennari, R. and Vittorini, P. (2010), “The design of an intelligent adaptive learning system for poor comprehenders”, *2010 AAAI Fall Symposium Series*.

Felten, E.W., Raj, M. and Seamans, R. (2023), “Occupational Heterogeneity in Exposure to Generative AI”, SSRN Scholarly Paper, Rochester, NY, 10 April, doi: 10.2139/ssrn.4414065.

Feng, S., Patel, S.S., Wan, H. and Joshi, S. (2021), “MultiDoc2Dial: Modeling Dialogues Grounded in Multiple Documents”, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing,* pp. 6162–6176, doi: 10.18653/v1/2021.emnlp-main.498.

Forbus, K.D., Riesbeck, C., Birnbaum, L., Livingston, K., Sharma, A. and Ureel, L. (2007), “Integrating natural language, knowledge representation and reasoning, and analogical processing to learn by reading”, *Proceedings of the National Conference on Artificial Intelligence*, Vol. 22, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, p. 1542.

Freed, E.M., Hamilton, S.T. and Long, D.L. (2017), “Comprehension in proficient readers: The nature of individual variation”, *Journal of Memory and Language*, Vol. 97, pp. 135–153, doi: 10.1016/j.jml.2017.07.008.

Gerven, F.P., Juhani E. Tuovinen, Huib Tabbers, Pascal W.M. Van. (2003), “Cognitive Load Measurement as a Means to Advance Cognitive Load Theory”, *Cognitive Load Theory,* Routledge.

Gläser, J. and Laudel, G. (2010), *Experteninterviews Und Qualitative Inhaltsanalyse,* Springer-Verlag.

Glim, S., Körner, A., Härtl, H. and Rummer, R. (2023), “Early ERP indices of gender-biased processing elicited by generic masculine role nouns and the feminine–masculine pair form”, *Brain and Language*, Vol. 242, p. 105290, doi: 10.1016/j.bandl.2023.105290.

Greco, C.M., Simeri, A., Tagarelli, A. and Zumpano, E. (2023), “Transformer-based language models for mental health issues: A survey”, *Pattern Recognition Letters,* Vol. 167, pp. 204–211, doi: 10.1016/j.patrec.2023.02.016.

Gregor, S. and Hevner, A.R. (2013), “Positioning and Presenting Design Science Research for Maximum Impact”, *Management Information Systems Quarterly,* Vol. 37 No. 2, pp. 337–355.

Gregor, S., Kruse, L. and Seidel, S. (2020), “Research Perspectives: The Anatomy of a Design Principle”, *Journal of the Association for Information Systems*, Vol. 21, pp. 1622–1652, doi: 10.17705/1jais.00649.

Gregory, R.W. and Muntermann, J. (2014), “Research Note—Heuristic Theorizing: Proactively Generating Design Theories”, *Information Systems Research*, INFORMS, Vol. 25 No. 3, pp. 639–653, doi: 10.1287/isre.2014.0533.

Gupta, S., Rawat, B.P.S. and Yu, H. (2020), "Conversational Machine Comprehension: a Literature Review", *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 2739–2753, doi: 10.18653/v1/2020.coling-main.247.

Head, A., Lo, K., Kang, D., Fok, R., Skjonsberg, S., Weld, D.S. and Hearst, M.A. (2021), "Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols", arXiv, 27 April, doi: 10.48550/arXiv.2009.14237.

Hevner, A. (2007), "A Three Cycle View of Design Science Research", *Scandinavian Journal of Information Systems*, Vol. 19 No. 2, pp. 87–92.

Howard, P.J., Gorzycki, M., Desa, G. and Allen, D.D. (2018), "Academic Reading: Comparing Students' and Faculty Perceptions of Its Value, Practice, and Pedagogy", *Journal of College Reading and Learning*, Vol. 48 No. 3, pp. 189–209, doi: 10.1080/10790195.2018.1472942.

Jáquez-Pérez, S.I. and Villa-Maciel, M. del C. (2021), "Adaptive Learning of Reading Comprehension: Innovation with Technologies for Middle Education", *Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21)*, Association for Computing Machinery, New York, NY, USA, pp. 477–481, doi: 10.1145/3486011.3486497.

Kelloway, E.K. and Barling, J. (2000), "Knowledge work as organizational behavior", *International Journal of Management Reviews*, Vol. 2 No. 3, pp. 287–304, doi: 10.1111/1468-2370.00042.

Kendeou, P. and van den Broek, P. (2007), "The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts", *Memory & Cognition*, Vol. 35 No. 7, pp. 1567–1577, doi: 10.3758/BF03193491.

Kittur, A. (2017), "Towards a Universal Knowledge Accelerator", *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, Association for Computing Machinery, New York, NY, USA, p. 1, doi: 10.1145/3126594.3126668.

Kizilcec, R.F. (2016), "How Much Information? Effects of Transparency on Trust in an Algorithmic Interface", *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 2390–2395, doi: 10.1145/2858036.2858402.

Kohnke, L., Moorhouse, B.L. and Zou, D. (2023), "ChatGPT for Language Teaching and Learning", *RELC Journal*, SAGE Publications Ltd, pp. 1–14, doi: 10.1177/00336882231162868.

Kosch, T., Schmidt, A., Thanheiser, S. and Chuang, L.L. (2020), "One does not Simply RSVP: Mental Workload to Select Speed Reading Parameters using Electroencephalography", *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 1–13, doi: 10.1145/3313831.3376766.

Lee, B., Savisaari, O. and Oulasvirta, A. (2016), "Spotlights: Attention-Optimized Highlights for Skim Reading", *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 5203–5214, doi: 10.1145/2858036.2858299.

Lee, M., Srivastava, M., Hardy, A., Thickstun, J., Durmus, E., Paranjape, A., Gerard-Ursin, I., *et al.* (2022), "Evaluating Human-Language Model Interaction", arXiv, 20 December, doi: 10.48550/arXiv.2212.09746.

Liang, H., Liu, J., Du, W., jin, dingnan, Lei, W., Wen, Z. and Lv, J. (2023), "Knowing-how & Knowing-that: A New Task for Machine Reading Comprehension of User Manuals", arXiv, 7 June, doi: 10.48550/arXiv.2306.04187.

Liew, A. (2007), "Journal of Knowledge Management Practice", *Journal of Knowledge Management Practice*, Vol. 8 No. 2.

Lo, K., Chang, J.C., Head, A., Bragg, J., Zhang, A.X., Trier, C., Anastasiades, C., *et al.* (2023), "The Semantic Reader Project: Augmenting Scholarly Documents through AI-Powered Interactive Reading Interfaces", arXiv, 23 April.

Ma, C., Wu, Z., Wang, J., Xu, S., Wei, Y., Liu, Z., Guo, L., *et al.* (2023), "ImpressionGPT: An Iterative Optimizing Framework for Radiology Report Summarization with ChatGPT", arXiv, 17 April.

Mcquillan, J. and Conde, G. (1996), "The Conditions of Flow in Reading: Two Studies of Optimal Experience", *Reading Psychology*, Routledge, Vol. 17 No. 2, pp. 109–135, doi: 10.1080/0270271960170201.

Nadj, M., Rissler, R., Adam, M., Knierim, M.T., Li, M.X., Maedche, A. and Riedl, R. (2023), "What Disrupts Flow in Office Work? The Impact of Frequency and Relevance of IT-Mediated Interruptions", *MIS Quarterly*, No. Forthcoming.

Nakamura, J. and Csikszentmihalyi, M. (2009), "Flow theory and research", *Handbook of Positive Psychology*, Vol. 195, p. 206.

"OpenAI customer story: Viable". (n.d.). , available at: https://openai.com/customer-stories/viable (accessed 17 November 2023).

Paranjpe, D. (2009), "Learning document aboutness from implicit user feedback and document structure", *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, Association for Computing Machinery, New York, NY, USA, pp. 365–374, doi: 10.1145/1645953.1646002.

de-la-Peña, C. and Luque-Rojas, M.J. (2021), "Levels of Reading Comprehension in Higher Education: Systematic Review and Meta-Analysis", *Frontiers in Psychology*, Vol. 12, pp. 1–11, doi: 10.3389/fpsyg.2021.712901.

Porsdam Mann, S., Earp, B.D., Møller, N., Vynn, S. and Savulescu, J. (2023), "AUTOGEN: A Personalized Large Language Model for Academic Enhancement—Ethics and Proof of Principle", *The American Journal of Bioethics*, Taylor & Francis, Vol. 23 No. 10, pp. 28–41, doi: 10.1080/15265161.2023.2233356.

Schlagwein, D. and Willcocks, L. (2023), "'ChatGPT et al.': The ethics of using (generative) artificial intelligence in research and science", *Journal of Information Technology*, SAGE Publications Ltd, Vol. 38 No. 3, pp. 232–238, doi: 10.1177/02683962231200411.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017), "Critical analysis of Big Data challenges and analytical methods", *Journal of Business Research*, Vol. 70, pp. 263–286.

Soemer, A. and Schiefele, U. (2019), "Text difficulty, topic interest, and mind wandering during reading", *Learning and Instruction*, Vol. 61, pp. 12–22, doi: 10.1016/j.learninstruc.2018.12.006.

Thissen, B.A.K., Menninghaus, W. and Schlotz, W. (2018), "Measuring Optimal Reading Experiences: The Reading Flow Short Scale", *Frontiers in Psychology*, Vol. 9.

Towey, C.A. (2000), "Flow", *The Acquisitions Librarian*, Routledge, Vol. 13 No. 25, pp. 131–140.

Venable, J., Pries-Heje, J. and Baskerville, R. (2016), "FEDS: a Framework for Evaluation in Design Science Research", *European Journal of Information Systems*, Vol. 25 No. 1, pp. 77–89.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User Acceptance of Information Technology: Toward a Unified View", *MIS Quarterly*, Management Information Systems Research Center, University of Minnesota, Vol. 27 No. 3, pp. 425–478, doi: 10.2307/30036540.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R. and Cleven, A. (2015), "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research", *Communications of the Association for Information Systems*, Vol. 37, Article 9, pp. 205–224.

Wang, K.-L., Sun, T.-T. and Xu, R.-Y. (2023), "The impact of artificial intelligence on total factor productivity: empirical evidence from China's manufacturing enterprises", *Economic Change and Restructuring*, Vol. 56 No. 2, pp. 1113–1146, doi: 10.1007/s10644-022-09467-4.

Zhang, H., Liu, X. and Zhang, J. (2023), "Extractive Summarization via ChatGPT for Faithful Summary Generation", arXiv, 9 April.

Zhang, Z., Yang, J. and Zhao, H. (2021), "Retrospective reader for machine reading comprehension", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, pp. 14506–14514.