Harmonizing Assistance: Moderating Visual andTextual Aids in AI-Enhanced Textbook Readingwith IRead

Xiaoyu Zhang
xyzhang.cger@gmail.com

ETH Zurich

Vincent Dörig
ETH Zurich

Peng Cui
ETH Zurich

Vilém Zouhar
ETH Zurich

Torbjorn Netland
ETH Zurich

Mrinmaya Sachan
ETH Zurich

Research Article

Keywords: Textbook Reading Support, Learning Engagement, Question Generation

Posted Date: August 23rd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4770422/v1

Additional Declarations: No competing interests reported.

Harmonizing Assistance: Moderating Visual and Textual Aids in AI-Enhanced Textbook Reading with *IRead*

Xiaoyu Zhang[1,3*†], Vincent Dörig[2†], Peng Cui[2], Vilém Zouhar[2], Torbjörn Netland[3], Mrinmaya Sachan[2]

[1*]ETH AI Center, ETH Zürich, Zürich, 8092, Switzerland.
[2]Department of Computer Science, ETH Zürich, Zürich, 8092, Switzerland.
[3]Department of Management, Technology, and Economics, ETH Zürich, Zürich, 8006, Switzerland.

*Corresponding author(s). E-mail(s): xyzhang.cger@gmail.com;
Contributing authors: vdoerig@student.ethz.ch; peng.cui@inf.ethz.ch; vilem.zouhar@inf.ethz.ch; tnetland@ethz.ch; msachan@ethz.ch;
[†]These authors contributed equally to this work.

Abstract

Textbooks continue to be one of primary mediums of learning. Students often need additional support during the process of reading textbooks leading to several research efforts that aim to increase student engagement and provide tailored experiences in textbook reading. However, providing excessive information beyond the textbook can also distract students from the reading task. When enhancing the reading experience, one has to strike a delicate balance between providing sufficient informational support and maintaining students' focus on textbook reading. Fusing together latest developments in large language models (LLMs), their applications in education and several pedagogical theories, we design a textbook reading guidance mechanism. We introduce *IRead*, an interactive tool for textbook reading which uses LLMs with visualization and interaction techniques, to enhance students' reading and learning experiences. *IRead* incorporates conceptual visualizations that reflect the textbook's content and features an AI-driven question bot that generates questions and offers hints in response to student reading and interaction history. We evaluate *IRead* with a between-subject user study and measure the effectiveness of our methodology in supporting the students' reading experience based on the Bloom's Taxonomy and the ARCS model. We collect feedback from participants ranging from undergraduate to doctorate students. The results highlight the effectiveness of simple yet

intuitive visualizations, such as the concept tree in *IRead*. We also derive general insights for the development of tools that enhance educational reading experiences.

Keywords: Textbook Reading Support, Learning Engagement, Question Generation

1 Introduction

With the latest advances in large language models (LLMs) and the increased focus on Human-AI collaboration, traditional methods of learning and teaching are undergoing significant transformations, introducing both new opportunities and challenges to educational settings. The conventional process of textbook reading seems increasingly redundant given the capabilities of LLMs to provide high-quality summaries [1–3], retrieve information [4, 5], and answer questions [6, 7]. Institutions worldwide are looking for ways to integrate LLM in teaching and learning but are also concerned about risks for counterproductive outcomes [8–10]. However, they also voiced concerns that these new methods might completely replace traditional textbook reading practices, which would conflict with the university's objective of fostering comprehensive student development. Motivated by these insights, we explore and design a novel textbook reading tool for college students grounded in pedagogical principles, aiming to enhance rather than distract from the learning process.

The core challenge of our research lies in maintaining a balance between providing sufficient supplementary information and avoiding distracting readers with an excess of it. Many existing AI-enhanced textbook or document reading tools aim to leverage the full potential of AI to offer maximum support. In contrast, we position our work within the context of higher education, focusing on designing mechanisms to control the volume of visual and textual information and deliver it at the appropriate times for the users. For instance, consider two existing reading products: Apple Books and Kindle. Apple Books offers extensive reading materials with internal and external links, while Kindle provides a more streamlined reading environment. Interestingly, students tend to use Apple Books for leisure reading and Kindle for more serious academic reading [11]. This dichotomy underscores the challenge of balancing educational objectives with student engagement. From a pedagogical standpoint, our goal is to encourage students to engage in deeper reflection and discover the topic independently, yet many students prefer seeking the quickest and most straightforward path to complete their reading tasks [12, 13]. Therefore, determining the right moments and methods to provide the appropriate level of informational support to student readers is the primary research focus of our work.

Informed by principles of guided exploration learning [14–16] and constructivist learning theory [17], we developed the overarching strategy of this project to employ questions as a means to guide and stimulate the reading process for students. We first assessed the viability of this approach and our target users' preferences regarding the questions through a preliminary study involving fifteen college students. This study also enabled us to identify four primary types of interactions favored by participants: highlighting text, taking notes, checking concept definitions, and content searching. We then developed a specific question-triggering mechanism aligned with the ARCS (Attention, Relevance, Confidence and Satisfaction) model [18] and created an AI-assisted textbook reading tool, *IRead*, to implement this mechanism. To equip *IRead* with high-quality questions, we train a question generation model on a textbook corpus sourced from Openstax[1] [19] and adapt it to suit our interaction mechanism. Furthermore, we enhanced *IRead* with a table of contents minimap and a collapsible concept tree to offer students an overview and motivate learning from various perspectives.

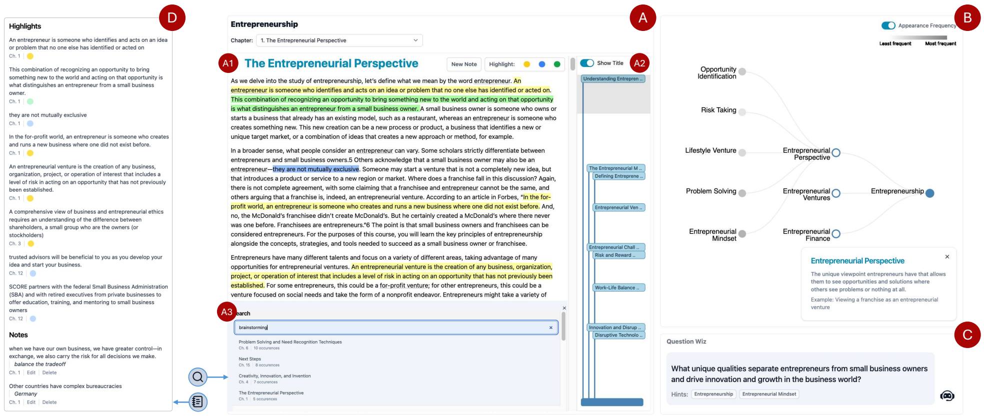


Fig. 1: The *IRead* interface comprises three major components: (1) A1 - an enhanced textbook reader with A2 - a minimap for logical overview and A3 - a query function for keyword searching; (2) B - a concept visualization for conceptual overview and interactive exploration of concept definitions; and (3) C - Question Wiz, which presents user interaction-specific questions for active and engaged reading.

Our framework and *IRead* were evaluated through a between-subject comparison study that adhered to the principles outlined in the ARCS model and Bloom's taxonomy [20]. Overall, the participants favor the novel reading experience provided by *IRead*, which also triggered more diverse thinking compared to the baseline system. The results also indicate that participants with different education levels, knowledge backgrounds, and reading habits demonstrated diverse preferences towards the components of *IRead*. From the study, we see the power of simple yet intuitive visualizations, like the concept tree in *IRead*, and have observed how various users employ it in diverse ways to serve specific purposes.

In summary, our major contributions are as follows:

1. A framework and mechanism designed to balance the trade-off between providing supplementary AI-generated information and minimizing distractions while learning by reading a textbook.
2. An enhanced textbook reading tool, *IRead*, that integrates the framework and mechanisms to facilitate a more engaging learning experience.
3. A preliminary study to assess user needs and preferences, followed by a comprehensive user study from various perspectives to gain insights and foster discussions on the effectiveness and user perception of the enhanced reading system.

[1] https://openstax.org/

2 Related Work

2.1 Guided Discovery Learning for Learning Engagement

The problem of driving learning engagement has existed in pedagogy for a long time and has been approached through different approaches such as using instructions [21] or questions associated with the text [22]. Among the existing attempts, one widely accepted theory is discovery learning developed by Jerome Bruner, John Dewey et al., where the student is provided with the learning materials rather than the exact answer so that they could find the answer themselves [23]. This learning technique also aligns with the constructivist-based education principles [17], where students actively construct their understanding and knowledge by integrating new information with their prior knowledge, rather than passively receiving new information. In this work, we facilitate discovery learning and enhance student engagement by developing an interactive textbook exploration tool where we generate questions regarding readers' reading history but not provide answers.

However, the cognitive load theory reveals that "*free exploration of a highly complex environment may overload working memory and hinder learning*" [24]. To balance between didactic teaching and unassisted discovery learning, researchers have developed the guided discovery learning approach [14–16] In practice, such a guided discovery learning experience can be achieved through inquiry-based learning [25], which fosters a deeper understanding of the material with reduced cognitive load. We combine teacher expertise with LLMs to develop a facilitator that guides students through their learning adventures. This approach encourages students to seek answers through further exploration of the textbook, resulting in iterative rounds of discovery and constructivist-based learning.

2.2 Reading Facilitate Techniques

Enhancing digital reading experience has long been explored by researchers in both pedagogy and computer science. Before the advent of AI, various methods were developed to individually improve the reading experience, commonly measured by reading comprehension performance. One of the simplest methods is note-taking, where readers independently take notes and summarize the texts they read, demonstrating significant benefits [26, 27]. Another popular technique is manual text highlighting, which has also been effectively implemented with digital tools to improve the reading experience [28].

In recent years, the surge of AI has created more possibilities for enhancing the reading experience. Natural language processing (NLP) techniques now enable automatic highlighting of extracted keywords or concepts and displaying their definitions for readers [1, 29]. Vision-language models can automatically enhance textbooks by matching images from the web with the textbook content [30]. Beyond AI solutions alone, studies have shown that instructors seek high-quality human-AI teaming approaches to improve their teaching efficiency [31, 32] and support students in reading [32]. In response to this, Ruan et al. developed BookBuddy, which converts reading materials into an interactive, conversation-based tutor leveraging chatbot technology [33]. VR and AR technologies have also been applied to create immersive visual interactions for children with storybooks [34]. However, most of these interactive technologies are geared towards entertainment reading for children. For more serious reading, the challenge of balancing engagement and learning performance remains. While there are many efforts to incorporate question-answer practice into the reading process [35–37], few explore the integration of additional visual and textual aids in an interactive textbook reading tool. In this paper, we further explore this problem and provide a solution to facilitate reading with both visualization and question-generation technologies.

2.3 NLP Techniques for Question Generation

NLP techniques have been widely used in supporting learning and teaching, for example, creating educational content [38], simulating student behaviors [39], offering adaptive feedback [40], etc. Among them, question generation, the technique to automatically create questions relevant to a given context, carries great potential to facilitate active learning [41]. Existing question generation approaches typically adopt language models, such as BART [42] and T5 [43], to create questions in an end-to-end manner. However, question generation systems in the NLP community mainly focus on the relevance and fluency of questions while overlooking their educational values [44]. As a result, recent efforts incorporate pedagogical principles into the system design and demonstrate their effectiveness in various scenarios, such as math problem solving [45] and language learning [46]. Besides the end-to-end models, [31, 32] emphasize the iterative process of question creation and collect a suite of NLP tools for Human-NLP-collaborative question generation, including text summarization, paraphrasing, and negation transformation. However, the questions of most existing work are designed for assessment or practice purposes. In contrast, the question generation module in *IRead* is trained to mimic "*what an instructor would ask students **during** the reading*", aiming to create an interactive reading experience.

3 Design Requirements

We began informing the system design with a preliminary study adopted from an existing work [19]. We focused on analyzing the reading strategies and enriching the reading experience by generating questions. Such questions are often used by good writers to make the readers engage more with the text. In the preliminary study, the participants read different articles with additional questions displayed in a side panel (see experiment interface in Appendix A). This study was done in-person on 16 participants with a think-aloud protocol paired with an exit questionnaire.

Despite judgment subjectivity, participants consistently prefer questions that are intriguing, informative, and stimulate thought and reflection. These questions offer insights or highlight specific details, but they must be challenging and motivate engagement with the text. Participants tend to dislike questions that are too easy or have immediate answers within the article. Half the participants believe a question's relevance depends on whether the following sentences contain the answer. Shorter questions tend to be less distracting, with their distractibility rating inversely proportional to perceived helpfulness.

Some participants experienced a "cold start" during the reading task, expressing difficulty in getting into the context of an article. Initial preferences lean towards easier and more closely relevant questions at the article's beginning. As they familiarize themselves with the context after reading a few paragraphs, their inclination shifts towards more intriguing and divergent questions. Furthermore, participants' standards for defining a good question change based on the reading scenario, e.g., first-time reading vs. content reviewing, learning scenarios vs. examination scenarios, and serious learning vs. casual reading.

Based on participants' responses, we synthesize the following requirements for a comprehensive reading assistance tool:

R1 **Overview:** The tool should provide a high-level overview of the textbook's content and structure. This overview can take the form of a textual summary detailing the article content or a graphical diagram illustrating the logical structure, with the latter being the preferred choice for the majority of participants.

R2 **Comprehension Support:** (Attention-A2.1, A2.2) The tool should provide enhanced reading support in comparison to a standard PDF reader. In addition to conventional features like highlighting, note-taking, and content searching, the tool should also provide supplementary information for terms that may impede the reading process. This additional information may include definitions, examples, external links, and illustrative figures.

R3 **Intelligent Guidance:** (Confidence-C3.2 Expectations) The tool should aptly navigate the student's reading process, tailoring guidance to their individual reading pace. An intelligent chatbot can be employed to offer assistance when readers face challenges such as a cold start, reading barriers, or reading fatigue. The questions posed by the chatbot should be intelligently positioned, with an appropriate frequency, and of suitable complexity, ensuring that they enhance the reading process rather than becoming a distraction.

R4 **Customized Information Inquire:** (Attention-A5.3) The readers should have control over the extent of additional information they wish to access, catering to their individual reading preferences across various scenarios, such as first-time reading, in-depth reading, serious reading, and casual reading. This includes the capability to control the visibility of the information outlined in **R1** to **R3**.

4 Implementation

The content presented in *IRead* is generated through the data processing workflow illustrated in Fig. 2. This workflow is a one-time setup for every newly imported textbook, which includes generating a structured overview of the main ideas (concept hierarchy), creating the table of contents, and integrating the questions from the model into the textbook.

4.1 Question Generation Model

The core idea behind *IRead* is a question-driven active reading mechanism. Our goal is to use questions as an interactive device to maintain readers' interest and attention, and to guide them through the reading process. However, existing question generationsystems are not adequate for our needs because:

- question generation models in the NLP community mainly focus on factual questions that lack educational value [32, 44].
- Most questions are generated for assessment purposes, and it remains unclear how to use them in an interactive reading scenario, such as *when* to present them to the reader.

To address these limitations, we have developed a question generation model based on the theories described in [19] considering three perspectives:

From the data perspective, we build our question generation model upon a textbook corpus collected from Openstax. The dataset consists of 3,593 questions from 621 textbook

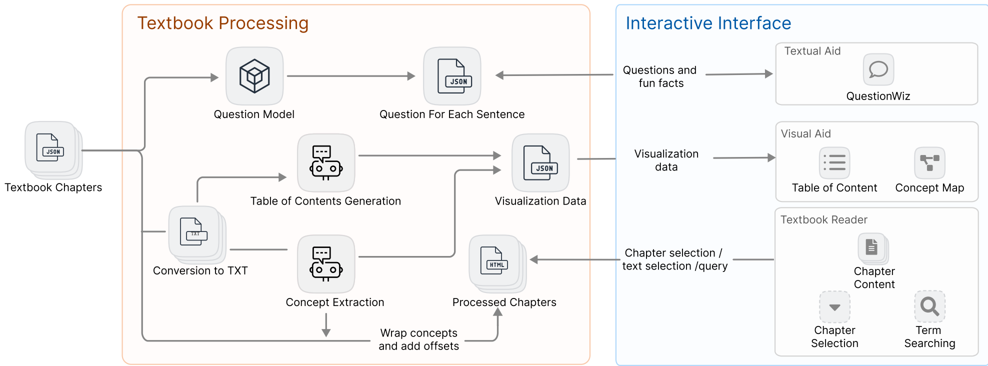


Fig. 2: Data processing and interaction workflow.

chapters across various disciplines, such as business and social sciences. In particular, instead of using assessment questions, we collect questions embedded in the chapter content. These in-text questions are designed by expert educators to provoke readers' thoughts *during* reading and learning. Meanwhile, they perform a variety of interactional functions to engage readers [47], therefore are the ideal choice for building our model.

From the model design perspective, our question generation model takes as input a passage and outputs a list of <*context*, *answer*, *question*>triples, where the *context* is the sentence that precedes the *question*, indicating at which point of the passage should the question be raised. For implementation, we fine-tuned the base version of Flan-T5 [48], an encoder-decoder-based language model. An illustration of the question generation model is available in Fig. 3.

From the interaction perspective, we modify the setup of the question generation model to adapt to our interaction mechanism (per Sec. 5.2). In particular, we enumerate every sentence as the context and generate the corresponding question. In doing so, we ensure that each sentence is associated with a relevant question that can be evoked upon the user's interaction. Building upon these base questions, we further employ ChatGPT to integrate user input and customize the questions displayed based on their specific interactions.

The question generation model was evaluated with both automatic and human evaluations, both of which demonstrated strong performance. For the automatic evaluation, we used reference-based metrics, achieving scores of 18.3 (Rouge-L [49]), 20.7 (Meteor [50]), and 84.5 (BertScore [51]), outperforming zero-shot GPT-4, which scored 15.7, 18.9, and 84.3, respectively. For the human evaluation, we invited 15 annotators to evaluate 32 model-generated questions from 5 textbook chapters, and another 15 annotators to evaluate 32 reference questions written by the textbook authors from the same chapters. The evaluation criteria included: relevance to the context, appropriate positioning without distraction, and importance to the central topic. On a scale of 1 to 5, the model-generated questions received scores of 4.2, 4.3, and 4.2 for the three criteria, respectively. These scores are comparable to those for the reference questions, which received 4.0, 4.2, and 4.4.

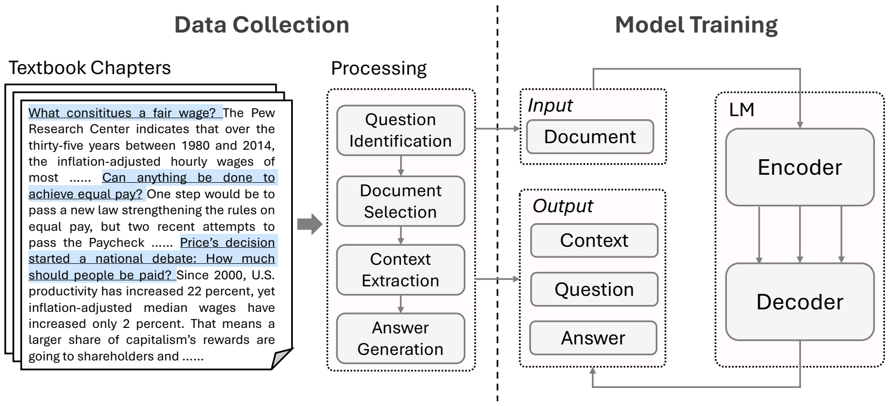


Fig. 3: Conceptual illustration of the question generation model.

4.2 Concept Tree Generation

In addition to questions, *IRead* includes a concept visualization feature to visually illustrate key concepts in the textbook and their hierarchical relationships. To generate the concept tree data for the visualization, we leverage the capabilities of LLMs, which excel in contextual comprehension and information synthesis. These models can also efficiently handle long texts due to their large context window sizes, allowing them to process an entire textbook chapter at once.

In practice, we instruct the LLM to identify key concepts and their related sub-concepts, along with their respective children, to structure the content hierarchically. We also request concept definitions and examples for inclusion in information cards and determine each concept's frequency within the chapter to illustrate its prevalence in the visualization. The resultant data is formatted as a nested JSON object.

We iterated through various versions of the prompt to identify one that consistently produces the desired data when applied to texts of varying lengths and domains. The prompt is provided in Appendix B-Listing 1, utilizing the GPT-4 (`gpt-4-turbo-preview`) model, which stands as OpenAI's latest and most advanced model at the time of development. After the auto-generation, we employ human oversight to ensure the balance of the tree structure and the quality of the concept, which are difficult to consistently manage through the prompt alone. Additionally, we offer an interactive visual interface that allows authors and instructors to further edit the concepts. This method of automated concept tree generation plus minimal post-inspection is efficient as it eliminates the need for additional inputs beyond the text itself while maintaining the concept tree quality. It is also cost-effective, as it only needs to be done once per chapter. Further evaluation and discussion on the quality of the concept tree are provided in Sec. 7.1.

4.3 Concept Mapping

To bridge the concept visualization with the textbook content, we implement a fuzzy concept mapping technique that aligns the concept tree with the text. Instead of restricting the mapping to exact matches, our approach aims to identify words or phrases closely related to the concepts, even with slight variations. To achieve this, we devise a method that utilizes synonyms generated by the LLM to pinpoint specific terms corresponding to each concept. We employ word embeddings that map words or sentences into a multi-dimensional space, using cosine similarity through a sentence-transformer model. This allows us to assess how well these terms correlate with the concept within a sentence. Terms exceeding a predefined similarity threshold are highlighted as clickable links and associated with a concept in the visualization within *IRead*. For a detailed implementation overview, refer to Algorithm 1.

Algorithm 1 Naïve Concept Mapping with Semantic Similarity

```
function MAPCONCEPTS(text, threshold)
    concepts ← EXTRACTCONCEPTSWITHLLM(text)
    sentences ← SPLITTEXT(text)                         ▷ split textbook content into sentences
    for all sentence ∈ sentences do
        for all word ∈ sentence do
            if word ∈ concepts then
                MARKASCONCEPT(word, concept)
                break
            end if
            w⃗ ← ENCODE(word, sentence)                  ▷ context-aware word embedding
            for all concept ∈ concepts do
                c⃗ ← ENCODE(concept, sentence)
                if (w⃗·c⃗)/(||w⃗||·||c⃗||) ≥ threshold then   ▷ cosine similarity threshold
                    MARKASCONCEPT(word, concept)
                end if
            end for
        end for
    end for
end function
```

5 *IRead*

To address the challenges discussed in Sec.1 and meet the design requirements outlined in Sec.3, we design a visual interface (Sec. 5.1) and an interaction mechanism (Sec. 5.2) grounded in multiple pedagogical theories, as illustrated in Fig. 4. The visual interface and interaction mechanisms are coordinated to create a cohesive, interactive learning environment, personalized for each user's exploration of textbook content.

5.1 Visual Interface

The *IRead* interface consists of three major components: an enhanced textbook reader (Fig. 1-A), a concept visualization(Fig. 1-B), and Question Wiz (Fig. 1-C). These components, along with their subcomponents, are designed to provide moderate visual and textual aids for users.

Enhanced Textbook Reader. To provide a fundamental reading experience similar to typical PDF readers like Acrobat, *IRead* includes annotative features to support active reading strategies. Users can highlight sentences in three different colors and add notes to selected text, which are automatically saved for future reference (**R2**). Furthermore, a sidebar (Fig.1-D), which can be displayed or hidden, organizes and displays the saved highlights and notes, allowing users to review and edit them as needed. To provide a linear overview and help readers navigate the text (**R1**), we include a minimap on the right side of the text, serving as a table of contents (Fig.1-A2). Clicking on the section titles takes the user directly to that section. Additionally, a search function allows users to perform queries across different chapters (Fig. 1-A3) and see how the same term is addressed in different contexts (**R4**).

Concept Visualization. Following the "overview first, details on demand" mantra, this component graphically represents the concepts from the textbook and their relationships within a hierarchical tree structure (**R1**). Concepts on the left side of the graph have higher-level and more abstract semantic meanings compared to those on the right side. When users hover over a concept, an information card appears with a definition and examples to aid comprehension (**R2**). Clicking on a node dynamically adjusts the tree to reveal the direct children of the selected concept. By expanding or collapsing the concept nodes, users can navigate through the tree, promoting an engaging and exploratory way of learning (**R4**). These concepts are also visually distinguished from other text with dotted underlines in the textbook reader. When a user clicks on such a term, the concept visualization updates to highlight this concept, revealing its information card and showing its family path in the tree. Meanwhile, Question Wiz will pose a relevant question, further engaging the user with the material.

Question Wiz. Located at the bottom right corner of the interface is the "Question Wiz", *IRead*'s reading assistant designed to show feedback from the various interaction mechanisms presented in Sec. 4 (**R3**). When a question is posed, the assistant not only presents the question but also offers hints in the form of related concepts to aid in formulating an answer. Users can interact with these hints by clicking on them, which highlights the respective concept in the visualization, strategically positioned right above to facilitate quick cross-referencing. Additionally, when a user moves to a new chapter via search, the assistant will offer to return them to the last reading position from the previous chapter.

Concept Mapping in the Text (Fig. 1-A1). In *IRead*, the user will see specific words visually distinguished from the others as dotted underlined. These denote concepts and clicking such a term triggers two response: it highlights the concept in the visualization and prompts the question bot to pose a relevant question, further engaging the user with the material. The concept mapping aids in comprehension and allows for cross-referencing between the textbook content and the visualization (**R2**).

5.2 Interaction Mechanism

Aligning with the most common reading practice, users usually begin their interaction with *IRead* by reading the textbook with the enhanced textbook reader. During this process, *IRead* supports highlighting and note-taking, similar to a conventional PDF reader, and stores these user-specific records in a history view for later review (aligning with constructivist learning principles). Moreover, users can query critical concepts they find challenging or search for keywords across multiple chapters in the textbook. *IRead* provides definitions and examples related to the concepts or presents the context of the keyword in all chapters for further understanding (aligning with inquiry-based learning principles). Additionally, when users highlight, take notes, search concepts, or pause reading with *IRead*, corresponding questions are triggered to encourage them to continue reading and explore concepts they are interested in (aligning with guided exploration learning principles).

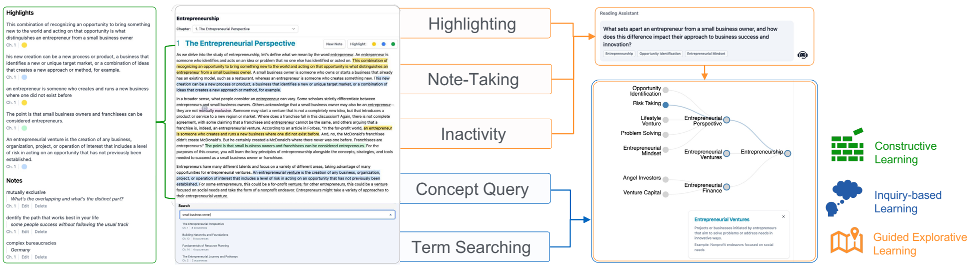


Fig. 4: The interaction mechanism of *IRead*. Informed by multiple pedagogical theories, we design the framework to trigger additional textual and visual support based on five different types of user interactions.

As mentioned above, the interaction-specific question generation mechanism is designed to provide readers with context-sensitive information based on their specific interactions with *IRead*. Drawing from insights gathered in the preliminary study, we implemented five types of interactions and aligned them with different question-triggering mechanisms according to the four dimensions—Attention, Relevance, Confidence, and Satisfaction—in Keller's ARCS model [18]. Table 1 presents the five interaction types, their triggering and question-display mechanisms, and the corresponding ARCS principles. We also assess the effectiveness of this mechanism through a questionnaire designed in accordance with the ARCS model, which we will discuss further in Sec. 7.3.3.

6 Use Case Scenario

6.1 Navigating New Knowledge as Novice Learner

In this use case, we describe a curious student with minimal background in psychology using *IRead* to read a psychology textbook [52] for a single-semester introductory psychology course. The student, interested in the topics of emotion and motivation, first browses the chapter titles and selects Chapter 10 to begin reading. Once the chapter is loaded, the student immediately notices a figure from the book illustrating intrinsic and extrinsic motivation. Curious about these terms, they click on the underlined concept *intrinsic motivation* from the figure title, which updates the visualization in the upper right corner and reveals the hierarchical position of the concept. There, they learn that intrinsic and extrinsic motivation are the two main types of motivation. They also hover over the concept to reveal both the definition and examples of these terms, grasping an overview of the conceptual grounding before delving into the reading material. Meanwhile, *IRead* simultaneously prompts a related question in the reading assistant panel as shown in Fig. 5-A. This question, along with hints to related concepts, helps the student engage with the material early on.

As the student progresses with the reading, they encounter a statement that aligns with their previous findings from the visualization:"*motivations can be intrinsic (arising from*

Interaction	Trigger	Question Wiz Display	Prompt for Customization	ARCS
Highlighting	Select and highlight a piece of text	A question related to the selected sentence with a list of relevant concepts below	*Paraphrase the following question such that it is intriguing, informative and stimulates thought and reflection:* {*original_question*}	Relevance
Note-Taking	Create a note for a selected piece of text	A question related to both the selected sentence and the note, with a list of relevant concepts below	*Generate a question regarding "*{*question*}*" based on the emotion/purpose indicated in* {*note*}*:*	Relevance
Concept Query	Click on an underlined concept in the reader	A question related to the concept with a list of hint concepts below	*Ask a question about the potential application of* {*concept*} *in the context of* {*chapter_name*}*. The question should be open-ended and thought-provoking. Return the question without any additional information.*	Attention Relevance
Term Searching	Search for a term, click on a search result to go to a different chapter and waits 5 seconds	A progress reminder of prior chapter reading, with a button to return to the prior chapter	-	Attention Relevance Satisfaction
Inactivity	There has been no activity on the page for two minutes	A fun fact about the current chapter	*Generate a fun fact for a college student reading a textbook called "*{*book_title*}*" in a chapter titled "*{*chapter_title*}*" and containing the following concepts:* {*concepts*}*. Simply return the fact without anything else.*	Attention

Table 1: Triggering operation, question-display mechanisms, and corresponding ARCS principles for the five primary types of interactions in *IRead*.

internal factors) or extrinsic (arising from external factors)". They choose to highlight this as an important point using the yellow highlighting color. Upon creating the highlight, a new question appears: "*What drives you to seek knowledge and pursue education from within yourself?*" Reflecting on and attempting to answer this question encourages a deeper personal connection to the content and enhances the student's learning experience.

Further into reading, the student notes the passage: "*physical reinforcement (such as money) and verbal reinforcement (such as praise) may affect an individual in very different ways*." Wanting to personalize this information, the student highlights it and creates a note stating, "*I am more likely to do something I dislike if I receive financial compensation*." They then pause to reflect on this topic and gradually drift away in thought. The system detects this inactivity and presents a fun fact about the current topic to re-engage the reader. This new information triggers the student's curiosity and motivation, prompting them to resume reading. Eventually, the student finishes the chapter with determination and enthusiasm, in an interactive and immersive reading environment fostered by *IRead*.

6.2 Reviewing Previous Content for Exam Preparation

The second use case illustrates how students can use *IRead* for reviewing and exam preparation. Consider a business student who used the system throughout the semester and is now leveraging it to reinforce and self-test critical concepts related to entrepreneurship in preparation for an exam. The student begins by carefully reviewing the concepts addressed in the chapter through the concept visualization, recalling the specific contexts in which they were

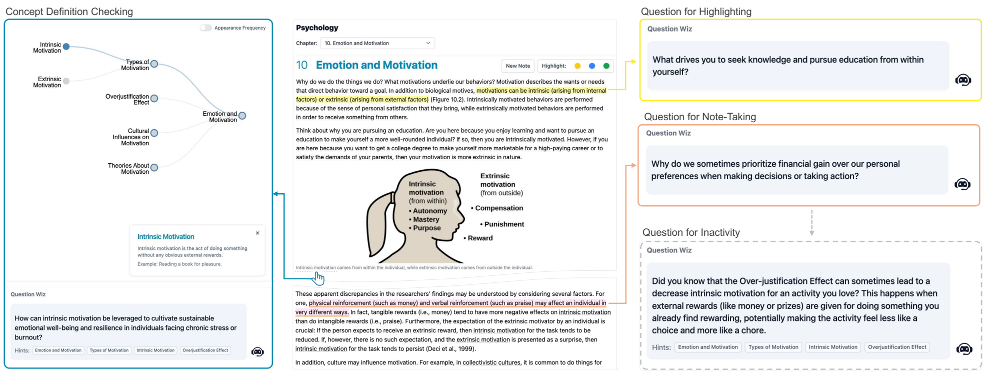


Fig. 5: *IRead* with a textbook on psychology while reading chapter 10 with "*intrinsic motivation*" as selected concept.

discussed. They also turn on the frequency toggle within the visualization to identify the key concepts to focus on. During this process, they realize they cannot clearly recall the meaning of "*entrepreneurial perspective*", so they hover over the concept to review its definition and click on it to explore the children concepts.

Afterwards, the student turns to the textbook content. Given that this textbook was utilized throughout the semester, it already contains numerous notes and highlights made by the user (Fig. 1-D), making it easy for them to navigate through the content. They then notice a notable new sentence: "*This combination of recognizing an opportunity to bring something new to the world and acting on that opportunity is what distinguishes an entrepreneur from a small business owner*" and highlight this sentence in green. This interaction triggers a question: "*What unique qualities separate entrepreneurs from small business owners and drive innovation and growth in the business world?*" Recognizing the importance of this question, the student spends extra time finding the answer. When faced with a similar question on the exam, they are able to answer correctly due to their thorough preparation.

As they continue reading, they come across another concept "*brainstorming*". Remembering that this concept is addressed in multiple chapters with different contexts, they use the search feature to locate additional information. The search bar provides a list of chapters where the concept occurs, along with the frequency of its mentions. The student navigates to the top chapter listed in the search results, "*Chapter6-Problem Solving and Need Recognition Techniques*", and review their previous highlights there (Fig. 1-A3). Recognizing this brief disruption, Question Wiz suggests returning to the prior reading and finishing reviewing the entire chapter by reminding the reading progress. The student accepts the suggestion and picks up their prior reading context. Finally, the student accesses all their notes and highlights in the sidebar, double-checking all concepts in the concept tree once more. Now, they feel well-prepared and confident for the upcoming exam.

7 Evaluation

7.1 Textbook Author Review

To understand the perspective of textbook authors on *IRead*, we showcased the tool using their own textbook and conducted casual interviews to gather their feedback. Below, we outline the key insights gleaned from these interviews.

AI-generated Content Quality. When inspecting the information provided by *IRead*, the book authors can quickly identify any discrepancies between their own knowledge and the content generated by the LLM. For instance, in the concept visualization, they believe the concept of ”manufacturing” has different meanings in the chapter, and therefore, it should be split into multiple subconcepts and distributed across lower levels of the concept map. Meanwhile, they acknowledge that AI cannot always distinguish such subtle differences or generate content that always satisfies the authors or instructors, but it provides a good starting point. Therefore, they suggest supporting more post-editing and customization based on the AI-generated content to ensure that “*at least we ensure nothing is wrong (before releasing it to the students)*”. They also recommend explicitly annotating which parts are generated by AI to avoid potential confusion.

Linear V.S. Non-Linear Reading. It is surprising to learn that the book authors have a fairly open attitude towards students' reading strategies. They do not intend to enforce the linear structure they designed for the textbook and agree that “*students have different reading strategies*”. One of the authors, who has experience instructing a master's course using the textbook they wrote, mentioned that they had tried to force students to read the textbook linearly, but “*it didn't work*”. Thus, they are willing to support students in learning non-linearly and constructively with *IRead*.

Novice V.S. Professional Reading Requirements. Although they are open to both linear and non-linear reading, the authors also suggest we support different reading preferences for novice and professional readers. Novice readers are usually more suitable for linear reading as they are unfamiliar with the organization of the content in the textbook. In contrast, professionals usually possess more domain knowledge and have specific information to find when they read. Their biggest challenge is usually “*not knowing exactly the name or where it is*”, so we need to facilitate a fast content location process. As evidenced in Sec. 7.3.3, the functionalities provided by *IRead* can already adeptly cater to these varied reading requirements.

7.2 Survey and Casual Interview in An Exhibit

We showcase *IRead* at an online exhibition and a half-day onsite Learning & Teaching Fair at a research university, attended primarily by professors, lecturers, scientists, and students from within and outside the university. During the fair, we demonstrated *IRead* through a poster and a short demo video to provide an overview. For those interested, we further provided a live demo to explain technical details, and invited attendees to complete a brief survey afterwards. We distributed the survey as hard copies and received 21 responses. The survey was designed to be completed within 10 minutes with a few 7-point Likert-scale questions regarding the participants' reading preferences and attitudes towards *IRead*.

The results show that 18 out of the 21 participants are open to reading the textbook in its digital version, and most are open to incorporating AI into their teaching and learning practices ($Avg = 6.10,\ IQR = 2$). Their ratings for *IRead* are positive ($Avg = 6.10,\ IQR = 1.5$), with the concept map and Question Wiz highlighted as the most useful and interesting features. Many participants expressed a willingness to try out the tool and stay updated on the latest developments of this work.

Interestingly, although participants have a positive impression of *IRead*, they appear less likely to integrate it into their daily teaching or learning routines. Many express concerns regarding the quality of the AI-generated content (5 mentions), its applicability to other documents such as academic papers or textbooks in math and computer science (2 mentions), and the need for real-world empirical studies with students (3 mentions). We believe the first two concerns can be alleviated with more demo time, allowing participants to explore our system more thoroughly. Addressing the last concern will be part of our future work, following comprehensive stability and scalability testing of the system.

7.3 User Study

We have designed and conducted a between-subject comparison study to explore our target users' preferences and assess the effectiveness of *IRead* in enhancing their motivation and learning dynamics during textbook reading. Through this user study, we aim to understand the following three research questions by gathering questionnaire responses and observing participants' behavioral patterns:

Q1: Can questions corresponding to various interactions effectively motivate users in reading and reflecting on the content? If so, what are the optimal content, frequency, and display timing for these questions?

Q2: Can the visualization we have designed aid in knowledge exploration? If so, what is the ideal level of detail and visual representation?

Q3: Can *IRead*, equipped with suitable prompted questions and visualization support, engage users in textbook reading and facilitate deeper content understanding?

In addition to these three research questions, we also aim to assess how well *IRead* fulfills the design requirements outlined in Section 3. Furthermore, we seek to gather insights into users' feedback on the usability of our system and potential enhancements.

7.3.1 Participants

We recruited 30 participants for the user study: 15 for the experiment group and 15 for the control group. All participants were current or past university students who read textbooks in their courses. They had good English reading and writing abilities and no mental, cognitive, or severe sight disabilities or impairments. On average, they read about 12 pages of text per day. The participants included 13 females and 17 males with educational backgrounds in computer science, management, biology, economics, and other fields. Among them were 2 undergraduate students, 16 Master's students, 11 doctoral students, and 1 postdoc researcher.

7.3.2 Setup and Procedure

To involve participants globally and simulate their typical learning environments, we offered the option for them to participate either remotely or in person. Tutorials for *IRead* were shared with participants at least 24 hours before the online study session began. We set up a remote server for remote participants to access *IRead* and join the study with their own machine and external devices. Participants were instructed to follow the "think aloud" protocol and their audio and screen activities were recorded during the task. The participants were rewarded with a voucher corresponding to $\sim$ \$27/hr.

To understand the effects of the visual and textual aids in *IRead*, we design the user study as a between-subject comparison study. Participants in the experiment group use *IRead* for the study, while those in the control group use a baseline variation with only the PDF reader and annotation functions (i.e., A1 panel in Fig. 1). Informed by two pilot studies, we structured the study with four parts: a tutorial session, two task sessions, and an experience survey and interview session. In the tutorial session, we demonstrated the usage of the reading tool and asked the participants to perform a few mini-tasks themselves. Afterwards, we ask them to perform below two tasks involving textbook reading and content question answering:

T1 **Textbook Reading.** Participants need to read the first half of a textbook chapter titled "The Entrepreneurial Perspective", available as an open-source resource on OpenStax. They are required to complete the reading within 25-30 minutes and were given the freedom to choose their preferred method of utilizing the assigned reading tool based on their individual reading habits.

T2 **Content Question Answering.** Participants answer six open-ended questions regarding the textbook content within a time frame of 15-20 minutes. We formulate the questions corresponding to the six levels of Bloom's taxonomy [20] to assess various depths of understanding of the participants. They are encouraged to refer to *IRead* as needed, simulating an open-book examination scenario.

After the two tasks, we ask the participants to respond to an experience survey adapted from the standard motivation questionnaires of ARCS [18, 53], which examines Attention, Relevance, Confidence, and Satisfaction individually. We also interview them promptly based on their responses to specific questions to better understand their learning experience, the reasons behind their actions, and the usability limitations of *IRead*.

7.3.3 Results

The statistical results show that *IRead* outperforms the baseline system in 8 out of 14 measured dimensions, with 5 achieving statistical significance (see Table 2). However, *IRead* tends to reduce participants' confidence in understanding the content and can sometimes be distracting. We provide further discussion and explanations based on observations and interviews from the study as follows.

Question is Not Just Question. Although the questions in *IRead* were originally intended to engage the reader, the study shows that participants use them for various purposes, including but not limited to previewing upcoming content, facilitating comprehension, and reviewing sections they have completed. For instance, P1 mentions that "*sometimes it takes the context of the paragraph that's coming next...so you know what's the main point of the paragraph, and so you can read it even faster*". P12 creatively utilizes the hints of the questions to "*pull up the concepts they want to check*". In addition, many participants find the questions regarding their highlighted content to be "*good as a recap*" (P3, P7) and "*trigger reflection*" (P4, P10, P14). As P3 commented, "*It satisfies in a way that you feel like you*

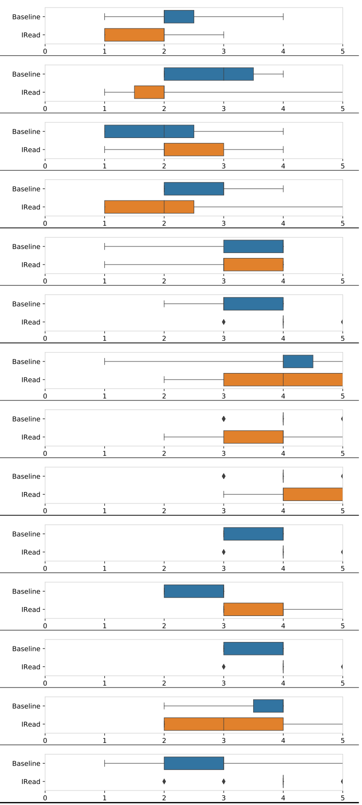
	Statement	Condition	N	M	SD	P	Agreement to the Statement (1 - 5)
Attention	**S1**: I feel rather disappointed with this tool.	Baseline	15	2.27	0.68	0.1413	
		IRead	15	**1.87**	0.72		
	S2: This tool has very little in it that captures my attention.	Baseline	15	2.93	0.77	**0.0086****	
		IRead	15	**2.0**	0.97		
	S3: I am often distracted while reading with this tool.	Baseline	15	2.0	0.89	0.2226	
		IRead	15	2.4	0.8		
	S4: I do NOT think I will benefit much from this tool.	Baseline	15	2.73	0.77	0.1081	
		IRead	15	**2.07**	1.29		
Relevance	**S5**: The reading support provided by this tool helps me address my confusion.	Baseline	15	3.4	0.95	0.3378	
		IRead	15	3.07	0.85		
	S6: The things I read with this tool will be useful to me.	Baseline	15	3.4	0.61	**0.0038****	
		IRead	15	**4.0**	0.37		
Confidence	**S7**: As I read with this tool, I believe I can understand all textbook content if I try hard.	Baseline	15	3.93	1.0	0.7476	
		IRead	15	3.8	1.17		
	S8: I feel confident that I will do well in the question answering section.	Baseline	15	3.87	0.5	0.4615	
		IRead	15	3.67	0.87		
	S9: I find the tasks of reading and answering questions to be manageable for me.	Baseline	15	3.93	0.44	0.096	
		IRead	15	**4.27**	0.57		
Satisfaction	**S10**: I enjoy reading with this tool.	Baseline	15	3.67	0.47	**0.0457***	
		IRead	15	**4.0**	0.37		
	S11: I feel that reading with this tool gives me a lot of satisfaction.	Baseline	15	2.67	0.47	**0.0001****	
		IRead	15	**3.67**	0.7		
	S12: I am curious about the textbook content I was reading.	Baseline	15	3.93	0.77	0.8247	
		IRead	15	3.87	0.81		
	S13: This tool presents the textbook content in a way that fulfills my expectation and goals.	Baseline	15	3.53	0.81	0.3339	
		IRead	15	3.2	0.98		
	S14: This tool presents various supplementary information in an engaging way.	Baseline	15	2.67	1.01	**0.0004****	
		IRead	15	**4.0**	0.73		

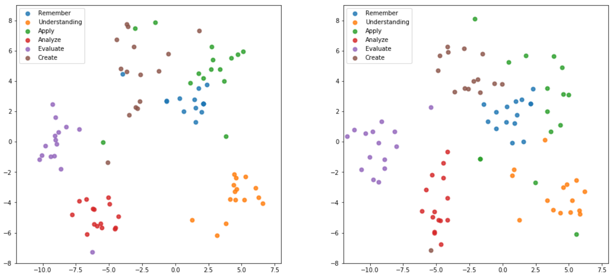
Table 2: Comparison between experiment group with IRead and a control group with the baseline PDF reader in IRead. Participants rated their agreement with questions related to the ARCS principles on a scale from 1 (strongly disagree) to 5 (strongly agree). Please note that the questions were reversed for *Attention*, so lower ratings were preferred. (M: mean, SD: standard deviation. * $p < 0.05$; ** $p < 0.01$)

learn double because it's not just reading a text ... I see where content is connected also with the question. It keeps me thinking about that and keeps me summarising. It's like, always a revision. It's not just reading words, and then you forgot. You continuously being in contact and remind of what you just read". In general, the prompt questions perform a positive role in inspiring readers' thoughts and improving learning effectiveness, which is also evidenced by the fact that participants in experiment group provides more variant answers to the content-related questions compared to the control group (Fig. 6).

Simple Visualization Can Be Powerful. *IRead* employs a simple concept tree visualization to present the semantic hierarchy of concepts. We have observed that the majority of participants react positively to this visualization and use it for various purposes in accordance with the design requirements. Their ratings for engagement and satisfaction are also

	Control Group	Experiment Group
Q1: Remember	0.37	**0.49**
Q2: Understanding	0.60	**0.62**
Q3: Apply	0.67	**0.76**
Q4: Analyze	0.54	**0.59**
Q5: Evaluate	0.61	**0.66**
Q6: Create	**0.72**	0.67

(a) Average In-Cluster Distance



(b) Control Group (c) Experiment Group

Fig. 6: The average in-cluster answer distance $d_{ans} = \frac{1}{n}\sum_{i=1}^{n} \|v_i - \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_j\|$ indicates that participants using *IRead* (experiment group) were more inspired in their thoughts and thus provided more varied answers compared to those using the baseline system (control group) for 5 out of the 6 content-related questions designed according to Bloom's Taxonomy. The 2D projections of the word embeddings of participants' answers also align with this insight.

significantly higher than those of the baseline group (see Table 2-S10, S11, S14). Four participants in the experiment group utilize the concept tree to gain an overview and navigate through "*the things they have to learn*" before reading (**R1**). Six participants use it more frequently during the reading process to check concept definitions (**R3**) and search for answers to questions indicated by question hints (**R2**). They comment that "*it's really clear...better to understand...even better than a definition from Google or Wikipedia*"(P1). Meanwhile, we have observed participants with a deeper background knowledge of the textbook topic (P5, P7) using the concept tree as a tool to review and confirm their understanding of all related key concepts in the chapter. In contrast, participants with almost zero background knowledge (P11) treat the concept tree itself as a learning material and directly learn from it. Additionally, we have observed three"*non-visual people*" (P3, P15) choose to skip the tree and use other functions to facilitate reading. This flexibility in using the same visualization in different ways demonstrates the level of customization we provide for users with varying educational backgrounds and personal reading habits (**R4**).

Question Timing is Critical. In the study, all participants in the experiment group comment and express their preferences regarding the timing of the questions. We track the timing of these questions based on interactions including highlighting, note-taking, and concept-definition checking initiated by the participants. Table 3 illustrates their opinions regarding questions about upcoming, current, and previously read content. We observed that most participants exhibited a consistent preference about upcoming and current content. This observation is also relevant to the convergent-thinking versus divergent-thinking reading habits mentioned in Section 7.3.3. When the type of questions they prefer dominates the reading process, participants tend to be more satisfied with the quality of the questions. It's also worth noting that P5's attitude towards all questions, especially early ones, changed from negative to positive after overcoming the learning curve and becoming accustomed to the questions in *IRead*.

Education Background Influence Preference Given that a significant number of participants find the distracting level of *IRead* to be high, we conduct a closer investigation and

Question Content	Upcoming	Current	Previous
Positive	4	4	0
Negative	5	4	1
Neutral	3	2	1
Not Trigger	3	5	13

Table 3: Statistics of participants' preference towards different questions timing. Questions about previously read content can only be triggered by highlighting a few heading sentences of subsections, so not all participants have the opportunity to encounter them.

discover that higher academic ranks (bachelor, master, doctoral, post-doctoral) correlate with lower perceived quality ($\rho = 0.2$) and increased reports of distraction ($\rho = 0.1$) (Pearson correlations with a significance level of $p < 5$ two-sided [54]) Participants in these higher academic ranks express a preference for receiving less visual and textual aid from *IRead* or no questions at all (see P12, P15). In contrast, participants with lower academic ranks described their ideal AI tools as providing "*intelligent summarization of content*" (P2, P10) and very high-quality questions (P10). This finding aligns with prior HCI studies that highlight the influence of user expertise [55, 56]. For different questions with objective, personal, or inspiring content, participants' opinions also vary. Divergent-thinking readers with business or management backgrounds (P3, P7, P14) prefer inspiring questions and questions that connect to their personal experiences. In contrast, convergent-thinking readers with mathematical or engineering backgrounds (P2, P13, P15) prefer more objective questions.

Reading Scenarios and State Influence Behavior Pattern. During the study, participants exhibit three distinct modes influenced by time pressure and task requirements. In the initial 15-20 minutes of **T1**, participants are relaxed to explore and comment on functions in *IRead*. They tend to utilize highlighting and note-taking functions more extensively and engage in question answering and visualization exploration to deepen their learning. As the reading task approached its 10-minute mark, we reminded participants about the time, participants were reminded of the time, prompting most of them (13 out of 15) to transition into a pressure mode and prioritize the completion of the reading task. In this mode, their usage was limited to highlighting and concept definition functions. Finally, during **T2**, simulating an open-examination scenario, participants utilize content searching functions more frequently for questions corresponding to the "memorize" and "understanding" levels of Bloom's taxonomy. Meanwhile, participants tend to review their own highlighting and note-taking history while also exploring the concept visualization for addressing higher-level open questions that require deeper personal reflection. Some participants noted that their behavior might vary when using *IRead* with textbooks from different disciplines. For instance, P3 mention their potential difference in behavior when using math textbooks, stating,"*I would have very specific questions... it's very hard to generate questions exactly focusing on this one...because you cannot make sure that it answered the questions I have, because maybe I don't understand an example.*" Moving forward, we plan to incorporate a customization component that involves instructors in contributing to the cultivation of the concept tree and the generation of questions.

7.4 Discussion

Information Support and Attention Guidance When using *IRead* to complete the reading task, four participants explicitly request that they want a Google Search function integrated in the tool to support searching of any selected keyword, and three mention they would like to directly communicate with the chatbot. These are what we intentionally avoided during the design stage, believing that this could easily distract the student readers to trivial topics. However, we learn from the study that this may also be relevant to the sense of security and confidence. For instance, P6 mentions that "*if I have more information, I feel safer. And then, I would like to learn more, not just the text*". They acknowledge that Google Search can be "*kind of distracting*", but they still feel is necessary. As a tradeoff, we plan to provide a deprived version of internal search, but also try to maintain the readers' attention on the by avoiding providing excessive information nor external links.

The Role of Visualization in AI Agent Reflecting on the long learning curve of many prior visual analytics systems, especially those with sophisticated workflow and AI-driven mechanisms [57, 58], we make a different attempt to employ a simple collapsible tree for the concept visualization. As demonstrated in Sec.7.3.3, this design has yielded overwhelmingly positive feedback and has met nearly all the requirements outlined in Sec.3. From this practice, we learned that visualization should prioritize reducing cognitive load rather than exacerbating it and convey information from AI agent to users in the most intuitive manner.

User Habits and Customization During the user study, we have also observed some very interesting usage habits and preferences. For example, both P3 and P7 prefer dense highlighting while reading. However, their opinions diverge on the generation of questions from closely highlighted sections. P3 believes that two highlights within the same sentence should yield different questions to prevent repetition, while P7 prefers consistency, as they typically make two consecutive highlights in concept definition sentences–one for the concept name and another for the detailed definition. Additionally, P5 expresses a preference for viewing all questions before reading to "*read with questions in mind*". In contrast, P7 and P15 prefer to read all questions after completing the reading to maintain uninterrupted focus. We aim to incorporate these diverse preferences into the next development iteration and further enhance the level of customization offered by *IRead* (**R4**).

8 Conclusion

In this paper, we present *IRead*, an interactive textbook reading tool grounded in a question-driven learning guidance mechanism developed through student feedback and multiple pedagogical theories. Our system integrates a state-of-the-art question generation model and an interactive visualization to enrich students' reading and learning experiences. Through surveys and interviews with textbook authors and university instructors, as well as a comparative study involving current or past students involved in higher education, we demonstrate the effectiveness of our approach in engaging students with textbook reading and the potential of applying human-AI collaboration technologies for educational purposes. Two universities have already expressed interest in integrating *IRead* into their teaching and learning systems, motivating our immediate next step to deploy and further evaluate its functionality in real-world teaching environments. Our ultimate goal is to make *IRead* publicly accessible and easily generalizable, allowing students of all locations, education levels, and knowledge backgrounds to benefit from the novel learning experiences offered by *IRead* and the new generation of AI-enhanced education.

Supplementary information. We have included a demo video to illustrate the interactions in *IRead* for the two use cases described in Sec. 6.

Acknowledgements. We thank the participants of our studies and the organizers and attendees of the Learning & Teaching Fair 2024 at ETH Zurich for their feedback to IRead. We are grateful to Mr. Michel Baudin for allowing us to test IRead on his textbook and for participating in our interviews as a textbook author. We extend our thanks to Prof. Dr. Menna El-Assady for her high-level suggestions on this work. We also thank Ms. Katalin Tesch for curating the demo video for this submission. Last but not least, we appreciate the valuable feedback and suggestions from the reviewers.

Declarations

Not applicable

Appendix A Experiment Interface for Preliminary Study

The preliminary study described in Sec. 3 was conducted using a textbook reading interface, as illustrated in Fig. A1.

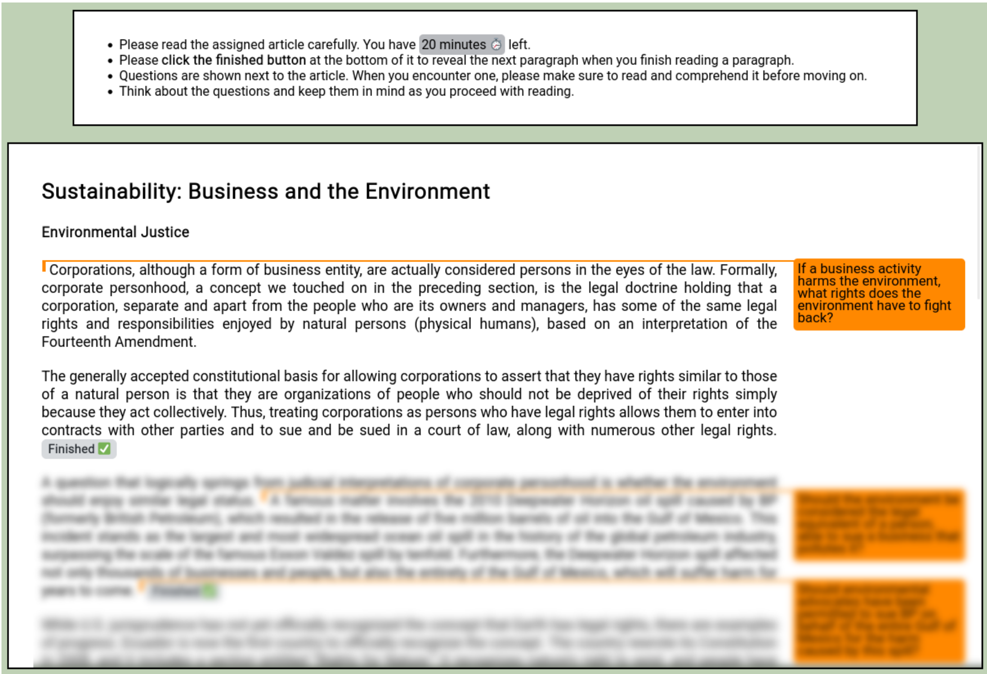


Fig. A1: The document contains generated active reading questions. Paragraph not yet been reached by the reader are blurred.

Appendix B LLM Prompts

Listing 1 is the prompt used to generate the concept tree data described in Sec. 4.2. Listing 2 is the prompt used to generate the table of content data mentioned in Sec. 5.1.

Listing 1: Concept Generation Prompt

```
Your task is to extract the primary overarching concepts given the text below.
     For each primary concept, discern a maximum of five sub-concepts with
    their own respective sub-concepts that are directly related. Ensure that
    the list of concepts does not become too long. Organize these concepts
    hierarchically, showcasing the main concepts at the top level and their
    corresponding sub-concepts indented beneath them. Aim to capture the
    fundamental structure of the document's concepts in a hierarchical format.
     The reader is a university student with good understanding in the text's
    field.

Return a nested JSON object with the following keys:
   - "name": the concept
   - "short_name": the abbreviated concept (if there is one or it is necessary
       )
   - "definition": a short definition of the concept
   - "example": an example of the concept (if applicable, may be a sentence or
        two)
   - a "verbatim_words" list: containing strings that appear in the text
       verbatim about the current concept (this list will be used to
       highlight the concept within the text to the user, be as specific as
       possible)
   - a "children" array: containing the children (if they exist) in the same
       format

Additionally, for the leaf concepts, add a "relative_frequency" key denoting
    the appearance frequency (between 0 and 1) and an "absolute_frequency"
    that is the absolute appearance count (number) of this concept in the text
    . For the absolute frequency, assign each sentence to one of the leaf
    concepts if applicable and sum them for each leaf concept. The relative
    frequencies should sum to 1.
```

Listing 2: Table of Contents Generation Prompt

```
Your task is to generate a table of content.

Return a nested JSON object with 3 levels. Every level includes following keys
    :
   - "index": section index as string
   - "title": the title of the current section (no more than 5 words)
   - "summary": a short summary of the current section
   - "length": the number of characters in the current section. It should
       equal to the total length of all its child subsections plus its own
       character count.
   - "char_start": the index of first character the corresponding section
```

```
  - "char_end": the lindex of last character of the current section. It
      should equal to char_end of its last children subsection.
  - "sen_start": the first sentence of the current section
  - "sen_end": the last sentence of the current section. It should equal to
      sen_end of its last children subsection.
  - "subsections": a list of subsections with the same keys as the current
      one
  - "depth": the depth number of current layer

Eliminate all empty object or list.
```

References

[1] Zhang, X., Li, J., Chi, P.-W., Chandrasegaran, S., Ma, K.-L.: ConceptEVA: Concept-based interactive exploration and customization of document summaries. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–16 (2023)

[2] Zhang, T., Ladhak, F., Durmus, E., Liang, P., McKeown, K., Hashimoto, T.B.: Benchmarking large language models for news summarization. Transactions of the Association for Computational Linguistics **12**, 39–57 (2024)

[3] Cohan, A., Dernoncourt, F., Kim, D.S., Bui, T., Kim, S., Chang, W., Goharian, N.: A discourse-aware attention model for abstractive summarization of long documents. In: Walker, M., Ji, H., Stent, A. (eds.) Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), pp. 615–621. Association for Computational Linguistics, New Orleans, Louisiana (2018). https://doi.org/10.18653/v1/N18-2097 . https://aclanthology.org/N18-2097

[4] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., *et al.*: Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems **33**, 9459–9474 (2020)

[5] Zhu, Y., Yuan, H., Wang, S., Liu, J., Liu, W., Deng, C., Chen, H., Dou, Z., Wen, J.-R.: Large language models for information retrieval: A survey. CoRR **abs/2308.07107** (2023) 2308.07107

[6] Ko, W.-J., Dalton, C., Simmons, M., Fisher, E., Durrett, G., Li, J.J.: Discourse comprehension: A question answering framework to represent sentence connections. In: Goldberg, Y., Kozareva, Z., Zhang, Y. (eds.) Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, pp. 11752–11764. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (2022). https://doi.org/10.18653/v1/2022.emnlp-main.806 . https://aclanthology.org/2022.emnlp-main.806

[7] Guu, K., Lee, K., Tung, Z., Pasupat, P., Chang, M.: Retrieval augmented language model

pre-training. In: International Conference on Machine Learning, pp. 3929–3938 (2020). PMLR

[8] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., *et al.*: Chatgpt for good? on opportunities and challenges of large language models for education. Learning and individual differences **103**, 102274 (2023)

[9] Jeon, J., Lee, S.: Large language models in education: A focus on the complementary relationship between human teachers and chatgpt. Education and Information Technologies **28**(12), 15873–15892 (2023)

[10] Tlili, A., Shehata, B., Adarkwah, M.A., Bozkurt, A., Hickey, D.T., Huang, R., Agyemang, B.: What if the devil is my guardian angel: Chatgpt as a case study of using chatbots in education. Smart learning environments **10**(1), 15 (2023)

[11] Moore, L.M.: At your leisure: assessing ebook reader functionality and interactivity. Unpublished dissertation). University College London, London (2009)

[12] Bharuthram, S.: The reading habits and practices of undergraduate students at a higher education institution in south africa: A case study. The Independent Journal of Teaching and Learning **12**(1), 50–62 (2017)

[13] Mizrachi, D., Salaz, A.M., Kurbanoglu, S., Boustany, J., Group, A.R.: Academic reading format preferences and behaviors among university students worldwide: A comparative survey analysis. PloS one **13**(5), 0197444 (2018)

[14] Brown, A.L.: Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. The journal of the learning sciences **2**(2), 141–178 (1992)

[15] Brown, A.L., Campione, J.C.: Guided Discovery in a Community of Learners. The MIT Press, Cambridge, MA (1994). doi.org/10.7551/mitpress/1861.001.0001

[16] Palincsar, A.S., *et al.*: Collaborative research and development of reciprocal teaching. Educational leadership **46**(4), 37–40 (1989)

[17] Matthews, M.R.: Constructivism in Science Education: A Philosophical Examination. Springer, Berlin, Germany (1998)

[18] Keller, J.M.: Development and use of the ARCS model of instructional design. Journal of instructional development **10**(3), 2–10 (1987)

[19] Cui, P., Zouhar, V., Zhang, X., Sachan, M.: How to Engage Your Readers? Generating Guiding Questions to Promote Active Reading (2024). https://arxiv.org/abs/2407.14309

[20] Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R., *et al.*: Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1:

Cognitive Domain. Longman, New York, NY (1956)

[21] Guthrie, J.T., Wigfield, A.: Reading Engagement: Motivating Readers Through Integrated Instruction. ERIC, Washington, DC (1997). https://eric.ed.gov/?id=ED404627

[22] Guthrie, J.T., Alao, S., Rinehart, J.M.: Literacy issues in focus: Engagement in reading for young adolescents. Journal of Adolescent & Adult Literacy **40**(6), 438–446 (1997)

[23] Dewey, J., *et al.*: The university elementary school: History and character. University Record **2**, 72–5 (1897)

[24] Kirschner, P.A., Sweller, J., Clark, R.E.: Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational psychologist **41**(2), 75–86 (2006)

[25] Bruner, J.S.: The act of discovery. Harvard educational review (1961)

[26] Rahmani, M., Sadeghi, K.: Effects of note-taking training on reading comprehension and recall. Reading **11**(2), 116–128 (2011)

[27] Özçakmak, H.: Impact of note taking during reading and during listening on comprehension. Educational Research and Reviews **14**(16), 580–589 (2019)

[28] Leroy, C., Kammerer, Y.: Reading multiple documents on a health-related issue: The roles of a text-highlighting tool and re-reading behaviour in integrated understanding. Behaviour & Information Technology **42**(14), 2331–2352 (2023)

[29] Zhang, X., Chandrasegaran, S., Ma, K.-L.: Conceptscope: Organizing and visualizing knowledge in documents based on domain ontology. In: Proceedings of the 2021 Chi Conference on Human Factors in Computing Systems, pp. 1–13 (2021)

[30] Singh, J., Zouhar, V., Sachan, M.: Enhancing textbooks with visuals from the web for improved learning. arXiv preprint arXiv:2304.08931 (2023)

[31] Lu, X., Fan, S., Houghton, J., Wang, L., Wang, X.: Readingquizmaker: A Human-NLP collaborative system that supports instructors to design high-quality reading quiz questions. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–18 (2023). https://dl.acm.org/doi/full/10.1145/3544548.3580957

[32] Wang, X., Fan, S., Houghton, J., Wang, L.: Towards process-oriented, modular, and versatile question generation that meets educational needs. In: Carpuat, M., Marneffe, M.-C., Meza Ruiz, I.V. (eds.) Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 291–302. Association for Computational Linguistics, Seattle, United States (2022). https://doi.org/10.18653/v1/2022.naacl-main.22 . https://aclanthology.org/2022.naacl-main.22

[33] Ruan, S., Willis, A., Xu, Q., Davis, G.M., Jiang, L., Brunskill, E., Landay, J.A.: Bookbuddy: Turning digital materials into interactive foreign language lessons through a voice chatbot. In: Proceedings of the Sixth (2019) ACM Conference on Learning@ Scale, pp. 1–4 (2019)

[34] Wang, Y., Mao, Y., Ni, S.-t.: Metabook: An automatically generated augmented reality storybook interaction system to improve children's engagement in storytelling. arXiv preprint arXiv:2405.13701 (2024)

[35] Johnson, B.G., Van Campenhout, R., Jerome, B., Castro, M.F., Bistolfi, R., Dittel, J.S.: Automatic question generation for spanish textbooks: Evaluating spanish questions generated with the parallel construction method. International Journal of Artificial Intelligence in Education, 1–20 (2024)

[36] Willis, A., Davis, G., Ruan, S., Manoharan, L., Landay, J., Brunskill, E.: Key phrase extraction for generating educational question-answer pairs. In: Proceedings of the Sixth (2019) ACM Conference on Learning@ Scale, pp. 1–10 (2019)

[37] Ruan, S., Jiang, L., Xu, J., Tham, B.J.-K., Qiu, Z., Zhu, Y., Murnane, E.L., Brunskill, E., Landay, J.A.: Quizbot: A dialogue-based adaptive learning system for factual knowledge. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)

[38] Sarsa, S., Denny, P., Hellas, A., Leinonen, J.: Automatic generation of programming exercises and code explanations using large language models. In: Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume 1. ICER '22, pp. 27–43. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3501385.3543957 . https://doi.org/10.1145/3501385.3543957

[39] Markel, J.M., Opferman, S.G., Landay, J.A., Piech, C.: GPTeach: Interactive TA training with GPT-based students. In: Proceedings of the Tenth ACM Conference on Learning @ Scale. L@S '23, pp. 226–236. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3573051.3593393 . https://doi.org/10.1145/3573051.3593393

[40] Liu, H., Liu, Z., Wu, Z., Tang, J.: Personalized multimodal feedback generation in education. In: Scott, D., Bel, N., Zong, C. (eds.) Proceedings of the 28th International Conference on Computational Linguistics, pp. 1826–1840. International Committee on Computational Linguistics, Barcelona, Spain (Online) (2020). https://doi.org/10.18653/v1/2020.coling-main.166 . https://aclanthology.org/2020.coling-main.166

[41] Kurdi, G., Leo, J., Parsia, B., Sattler, U., Al-Emari, S.: A systematic review of automatic question generation for educational purposes. International Journal of Artificial Intelligence in Education **30**, 121–204 (2020)

[42] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., Zettlemoyer, L.: BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In: Jurafsky, D., Chai, J., Schluter, N., Tetreault, J. (eds.) Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 7871–7880. Association for Computational Linguistics, Online (2020). https://doi.org/10.18653/v1/2020.acl-main.703 . https://aclanthology.org/2020.acl-main.703

[43] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P.J.: Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of machine learning research **21**(140), 1–67 (2020)

[44] Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., Yu, M., Wang, D., Li, T.J.-J.: Storybuddy: A Human-AI collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. CHI '22. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491102.3517479 . https://doi.org/10.1145/3491102.3517479

[45] Jiao, Y., Shridhar, K., Cui, P., Zhou, W., Sachan, M.: Automatic educational question generation with difficulty level controls. In: International Conference on Artificial Intelligence in Education, pp. 476–488 (2023). Springer

[46] Cui, P., Sachan, M.: Adaptive and personalized exercise generation for online language learning. In: Rogers, A., Boyd-Graber, J., Okazaki, N. (eds.) Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 10184–10198. Association for Computational Linguistics, Toronto, Canada (2023). https://doi.org/10.18653/v1/2023.acl-long.567 . https://aclanthology.org/2023.acl-long.567

[47] Hyland, K.: What do they mean? questions in academic writing. Text & Talk **22**(4), 529–557 (2002)

[48] Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., et al.: Scaling instruction-finetuned language models. arXiv preprint arXiv:2210.11416 (2022)

[49] Lin, C.-Y.: ROUGE: A package for automatic evaluation of summaries. In: Text Summarization Branches Out, pp. 74–81. Association for Computational Linguistics, Barcelona, Spain (2004). https://aclanthology.org/W04-1013

[50] Banerjee, S., Lavie, A.: METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In: Goldstein, J., Lavie, A., Lin, C.-Y., Voss, C. (eds.) Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation And/or Summarization, pp. 65–72. Association for Computational Linguistics, Ann Arbor, Michigan (2005). https://aclanthology.org/W05-0909

[51] Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., Artzi, Y.: Bertscore: Evaluating text generation with bert. arXiv preprint arXiv:1904.09675 (2019)

[52] Spielman, R.M., Jenkins, W.J., Lovett, M.D.: Psychology 2e. OpenStax, Houston, TX (2020)

[53] Ma, L., Lee, C.S.: Evaluating the effectiveness of blended learning using the ARCS model. Journal of computer assisted learning **37**(5), 1397–1408 (2021)

[54] Kowalski, C.J.: On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. Journal of the Royal Statistical Society: Series C (Applied Statistics) **21**(1), 1–12 (1972)

[55] Sauer, J., Seibel, K., Rüttinger, B.: The influence of user expertise and prototype fidelity in usability tests. Applied Ergonomics **41**(1), 130–140 (2010) https://doi.org/10.1016/j.apergo.2009.06.003

[56] Bubalo, N., Honold, F., Schüssel, F., Weber, M., Huckauf, A.: User expertise in multimodal HCI. In: Proceedings of the European Conference on Cognitive Ergonomics, pp. 1–6 (2016). https://dl.acm.org/doi/abs/10.1145/2970930.2970941

[57] Zhang, X., Xuan, X., Dima, A., Sexton, T., Ma, K.-L.: Labelvizier: Interactive validation and relabeling for technical text annotations. In: 2023 IEEE 16th Pacific Visualization Symposium (PacificVis), pp. 167–176 (2023). IEEE

[58] Zytek, A., Liu, D., Vaithianathan, R., Veeramachaneni, K.: Sibyl: Understanding and addressing the usability challenges of machine learning in high-stakes decision making. IEEE Transactions on Visualization and Computer Graphics **28**(1), 1161–1171 (2021)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- IReadV2.0.mp4