



Information Research – Vol. 31 No. ISIC (2026)

iSchool students' mental models of GenAI

Irene Lopatovska, Conor Mack, and Ellen Connors

DOI: <https://doi.org/10.47989/ir31ISIC65224>

Abstract

Introduction. Knowledge of users' mental models (an understanding of a system and its functions) helps align system behaviour with user expectations and identify user concerns that influence adoption. We examined graduate students' mental models of generative artificial intelligence (GenAI) as part of a larger study of GenAI adoption for academic purposes at the Pratt Institute iSchool.

Method. The data about mental models of GenAI were collected through drawings and conversations with participants during focus groups.

Analysis. The qualitative visual and textual data were analysed using thematic analysis.

Results. Students' mental models reflect a high-level understanding of the technical complexity of the GenAI system, with occasional acknowledgments of end-users and collection creators, as well as ethical, environmental, and quality concerns. mental models often reflect knowledge of prior information technology and use human analogies to explain the GenAI.

Conclusions. Even a technologically literate group of information science graduate students exhibits limitations in understanding GenAI systems. iSchool students are not only avid users of GenAI—whose deeper understanding of this technology could enhance usage efficiency and satisfaction—but also future designers of GenAI systems. Technology features, additional experiences with technology, educational resources/curricula, and critical media coverage could improve their understanding of GenAI's technological and social dimensions.

Introduction

Generative artificial intelligence, or GenAI, is a multimodal and multifunctional system that creates content, including text, images, audio, and code, by learning patterns in existing data through advanced techniques like deep learning and neural networks (Ronge et al., 2025). These systems are probabilistic, trained on large amounts of data to learn patterns and complex statistical relationships (Bandi et al., 2023), with outputs generated in small sections, or tokens, each being the most likely continuation based on statistical patterns formed from its training. GenAI outputs are non-deterministic, meaning the same input will not generate the same output twice (Bandi et al., 2023; Tankelevitch et al., 2024).

Since GenAI's introduction three years ago, it has seen widespread adoption, including by university students. Studies have shown that students use GenAI appropriately and inappropriately for academic tasks (Johnston et al., 2024; Digital Education Council, 2024; Kumar & Raman, 2022; Jin et al., 2025; McDonald et al., 2025; Reiter et al., 2025). While more studies examined GenAI uses by undergraduates, or younger (K-12) students' mental models of GenAI, fewer studies looked into university students' perceptions and mental models of GenAI. We focused on the graduate students' mental models of GenAI, and by engaging a sample of information science students, explored how technology-literate GenAI consumers and future designers understand this tool. Knowledge of students' mental models has the potential to align system behaviour with user expectations and identify user concerns that influence usage.

Literature Review

GenAI adoption by students

Research highlights the diversity of GenAI platforms used by students for various academic tasks, including brainstorming, writing, and programming (Digital Education Council, 2024). Recent reports suggest that students prefer general-purpose platforms like ChatGPT, and their usage depends on their academic needs. The majority of students view GenAI as a helpful tool for completing routine tasks and creating personalised educational approaches (Kumar & Raman, 2022; Maita et al., 2024). Students have been found to perceive GenAI as helpful (Amoozadeh, 2024) and easy to use (Daher and Hussein, 2024), though not always accurate or trustworthy. Wang and Li (2024) found that positive emotions, digital efficacy, and perceived usefulness affect students' willingness to adopt artificial intelligence (AI) tools for learning. However, when GenAI does not meet expectations, these attitudes can change (Amoozadeh, 2024) and/or lead to lower satisfaction and decreased intent to use (Daher & Hussein, 2024).

Studies illustrate that concerns about the fairness, ownership, accessibility, autonomy, privacy, security, societal impact, and sustainability affect GenAI usage (Baek et al., 2024; Hagendorff, 2020; Huang et al., 2023; Perkins, 2023; Maita et al., 2024; Siau & Wang, 2020; Zhang et al., 2021). Regardless of the levels of understanding of ethical issues, Hsiao and Tang (2024) demonstrated that it has an insignificant effect on GenAI usage.

Students' demographic characteristics, academic majors, exposure to, and awareness of GenAI have also been found to play an important role in the adoption process (Baek et al., 2024; Kelly et al., 2024; Johnston et al., 2024). Baek et al. (2024) found that age, gender, academic major, and institution type are significant factors influencing the use of GenAI. For example, students in their 30s and 40s, as well as male students, reported higher usage than participants in other age and gender groups. Daher and Hussein (2024) found that while male students perceived GenAI more positively than female students in terms of efficiency, affect, and intention to use, both groups reported similar levels of interaction.

Mental models of GenAI

Mental model (MM), defined as one's understanding of a system's functionality, is a well-established concept in cognitive science (Gero et al., 2020; Cannon-Bowers et al., 1993; Carroll & Reitman, 1988). Users of information systems rely on mental models to understand the system and its practical use, draw inferences, make predictions about its behaviour, and devise ways to interact with it.

Recent research focused on clarifying the characteristics of users' mental models of AI systems and identifying factors that influence the development of such models. Gero et al. (2020) have found that people's mental models include features related to a) global behaviour, AI performance across many turns (e.g., biases, typical mistakes); b) local behaviour, AI behaviour in a particular interaction; and c) knowledge distribution, what information the AI *knows* or can deduce and what it does not. The authors suggest that the accuracy of the MM depends not only on the understanding of the technical aspects of AI, but also on observations and interpretations of AI behaviour. AI's anthropomorphism has also been shown to affect mental models. For example, Doyle et al. (2021) discuss how interactions with conversational AI lead to the perceptions of AI as partners; Wang (2024) relies on the theory of mind, attribution of beliefs, desires, intentions, and emotions to oneself and others, to explain how the style of AI behaviour and its humanlike qualities influence people's mental models. DeVito (2021) applied the concept of folk theories, informal socially shared beliefs about systems' inner workings, to explain how users construct mental models about the invisible parts of GenAI (algorithms, ranking, moderation), drawing from experience, observation, community discussion, and platform changes. The study suggests that understanding of the system complexity and users' expectations leads to more detailed or sophisticated folk theories of AI. Additional factors that have been found to influence mental models include frequency of interactions and limitations of human cognition (Bansal et al., 2019), cognitive biases (DeVito, 2021; Nourani et al., 2021), individual differences in technological literacy (Druga et al., 2022; Long & Magerko, 2020), media coverage (Chen et al., 2025; Banks, 2020) and the degree of AI explainability (Ehsan et al., 2022; Chiyah Garcia et al., 2018; Liao et al., 2020; Rutjes et al., 2019). Development of accurate mental models of GenAI is hindered by systems' often unpredictable behaviours and outputs, and its performance adjustments based on the user-generated information and the system's 'mental models' of users (Wang et al., 2021; Ashktorab et al., 2019).

Students' mental models of GenAI

While the research on the university students' mental models of GenAI is scarce, several studies examined the mental models of K-12 students. Marx et al. (2023) reviewed 18 studies related to the secondary school students' (grades 6-12) attitudes and MM of AI. The authors found several common themes across the studies, including: a) the tendency to understand AI simplistically, as robots, intelligent computers, or automation, and conflate AI with general computing or data processing; b) frequent incorrect assumptions about the AI, for example, as unbiased and universally accurate; c) mental models that often emphasise anthropomorphic, rather than algorithmic features. Students' mental models are shaped by technology curricula (exposure to AI and programming), media presentations, social and cultural environments (including discussions with family and peers), and children's developmental stages. Marrone et al. (2025) also investigated secondary school students and their perceptions of AI as a collaborative teammate. Students generally found the AI tool to be 'cool' and engaging, even when it made errors, but preferred it in supportive, not equal, social roles. Students' mental models were frequently shaped by media portrayals, science fiction, documentaries, and marketing, often resulting in unrealistic expectations of AI. Students' perceptions of GenAI as a collaborator were also examined by Kim and Cho (2025), who found these perceptions to be shaped by media, algorithmic literacy, and public attitudes. The authors explored the different ways of explaining AI through technical,

analogical, or behavioral descriptions. They found that behavioural or analogy-based explanations led to more concrete and nuanced mental models compared to only technical or no explanations.

Sanusi et al. (2024) explored how social, educational, or technological interventions influence students' and teachers' MM and interactions with AI systems. The authors found that both groups' initial mental models relied heavily on anthropomorphic or familiar analogies, prior familiarity with AI, domain knowledge, and individual differences. Participants whose mental models better aligned with actual system capabilities achieved better use outcomes and expressed greater trust in technology. Baez et al. (2023), who also examined students' and teachers' perceptions of AI, highlighted the importance of motivation and self-efficacy in developing positive perceptions and effective use of AI by both groups in a learning setting.

Zirenko et al. (2025) investigated how the instructional interventions influenced university students' mental models. The initial mental models, which were often fragmented or simplified, became more accurate with instruction, leading to improvements in the use of technology. However, similarly to Sanusi et al. (2024), the authors identified individual differences in students' mental models and performance, with a subset of students reverting to simpler or erroneous models under stress or novel tasks, indicating that their mental models are not stable. Katsantonis and Katsantonis (2024) explored how social science (non-STEM) university students perceive AI across cognitive, emotional, and behavioural dimensions, and examined how demographic and contextual factors influence these perceptions. The authors found that emotional responses to AI (e.g., enthusiasm, affective disposition) drive intention more strongly than cognitive beliefs. The authors recommend introducing AI literacy curricula, integrating AI topics across courses, and scaffolding students' emotional and cognitive comfort with AI tools.

Zhu et al. (2024) examined how AI ethical anxiety, risks, and awareness shaped university students' intention and use of GenAI systems. The study found that ethical awareness both motivates adoption and increases perceived risk, and that ethical anxiety deters the actual use, even if it does not affect intention directly.

In summary, studies of school and university students typically reveal simplifications and inaccuracies in the initial mental models of AI systems, which are often influenced by media and experiences with GenAI. Studies tend to emphasise the importance of AI education in improving students' mental models and the effective use of GenAI.

Method

With many of the prior studies focusing on the factors that affect MM and exploring (sources of) interventions that could increase their accuracy, we prioritised understanding the main components of the mental models of GenAI. We conducted a study to investigate GenAI adoption by students of the Pratt Institute (New York City) School of Information (iSchool). The study used a convenience sampling to collect data about students' uses and attitudes towards GenAI through an online questionnaire (N=30; *the discussion of the questionnaire findings can be found in Lopatovska et al., in press*). Additionally, two focus groups (N=9) were conducted to expand the questionnaire findings and investigate students' mental models of GenAI (<http://bit.ly/42BmXje>). Students from the two largest programs of the Pratt iSchool, library and information science, and user experience design, volunteered for the study. This paper presents the findings on students' mental models of GenAI, with additional study results included to provide context for the discussion.

The data about mental models were collected during the focus groups, which lasted approximately 75 minutes. During the warm-up segment of the focus groups, participants were asked to visually represent how they think GenAI works using colourful pens and paper. Focus group moderators asked participants to label the main objects in their drawings for clarity before facilitating

discussions about both their drawings and their understanding of GenAI. This activity was followed by a discussion on the use of GenAI for academic tasks and students' attitudes towards this technology. Sessions were recorded, and three researchers took detailed notes. Participants received \$20 gift cards. The study was approved by the institutional review board.

The drawings of GenAI and the recordings of participants' verbal explanations were anonymised and analysed by two researchers using inductive thematic analysis. The researchers applied Braun and Clarke's (2012) six phases of thematic analysis to familiarise and immerse themselves in the drawings and narrative, build codes, group them into themes, and identify broader families after multiple rounds of iteration and re-examination to unearth a user's implicit and explicit understandings of GenAI. The codebook can be found at <http://bit.ly/3W2gFpi>, with the summary findings presented below. Participants are referred to with the letter "D" followed by a randomly assigned number.

Findings

All study participants used at least one GenAI tool for academic tasks, primarily brainstorming and writing. The most frequently used platform was ChatGPT, followed by Grammarly, Gemini, and other platforms. More than half of the participants used advanced prompting techniques in their GenAI interactions, and over 90% were skeptical about the quality of the GenAI outputs (Lopatovska et al., in press). Four focus group participants described themselves as advanced, four as intermediate, and one as a beginner GenAI user.

The analysed drawings and corresponding narratives shared many similarities, which were grouped into three high-level families: a system, a human, and value judgments.

The visual elements of the mental models that represented GenAI as a system of interconnected elements were grouped under the System family. Elements related to human interaction (or lack thereof) with the GenAI system formed the Human family. The Value Judgments family was formed by elements that represented personal assessments of GenAI and its ethical, social, and environmental impacts (complete coding hierarchy with examples can be found at <http://bit.ly/3W2gFpi>.)

Theme 1: The system

All nine drawings included elements of the GenAI system, represented with various degrees of detail and accuracy.

1.1 Accurate understanding

1.1.1 Pattern-seeking or clustering information (6/9)

Six drawings represented GenAI as a system that seeks patterns, matches, or clusters information (see example in Figure 1).

Participants' verbal descriptions of their models reiterated this understanding:

'To me, AI is about grouping similar things. For example, if I say, 'I like food,' the AI connects 'like' and 'food.' My mental model of AI is like a graph'. D5

1.1.2 GenAI as probabilistic system (5/9)

Five drawings represented GenAI as a probabilistic system where outputs are generated predictions (Figure 2 example).

1.1.3 Presence of information collections (5/9)

Five participants included information collections as a necessary component for GenAI training and/or processing (Figure 3).

'I have an arrow that says collection. So, the data will be collected, and I have a drawing of a computer and some servers. The data is collected and stored'. D6

Three representations of collections emphasised the human-generated nature of collections' content:

'I drew art, writing, science, and research, and then added a little thief stealing all of it... Then the AI trains all night while the sorting continues'. D8

'data like images and text are collected and stored on servers, then processed through algorithms and sent back to the user'. D6

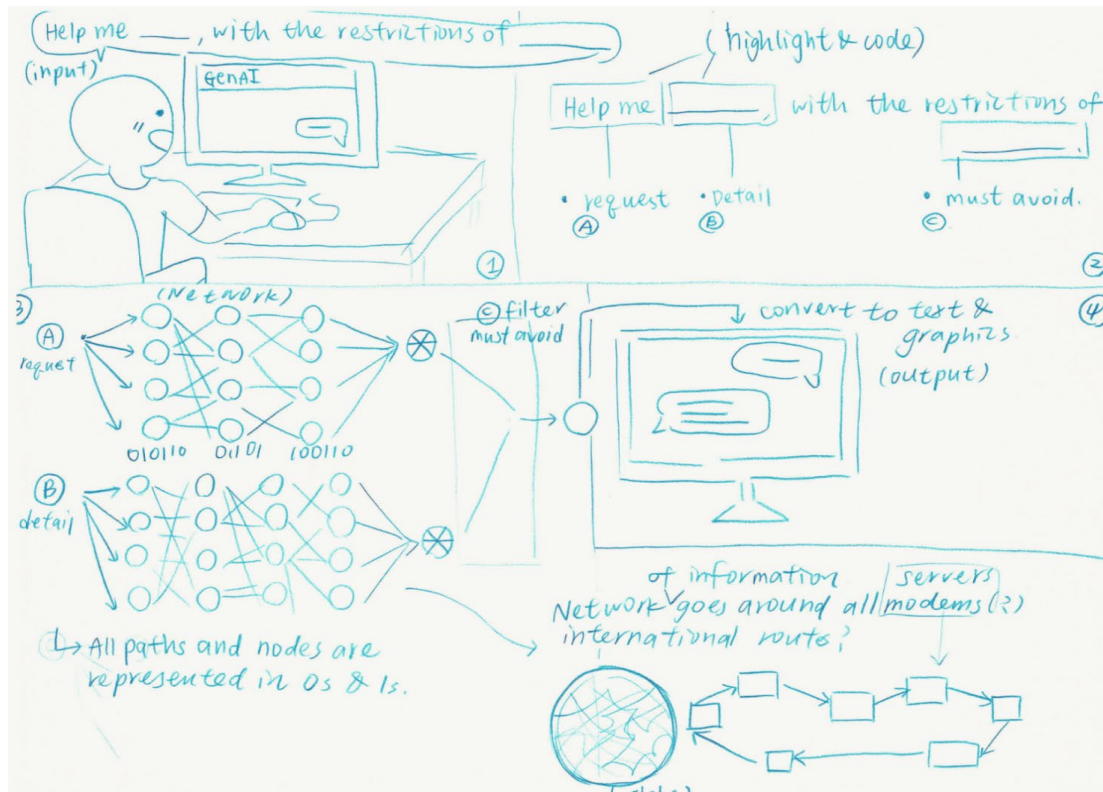


Figure 1. D7 MM illustrates clustering by analysing the structure of a prompt and generating its output.

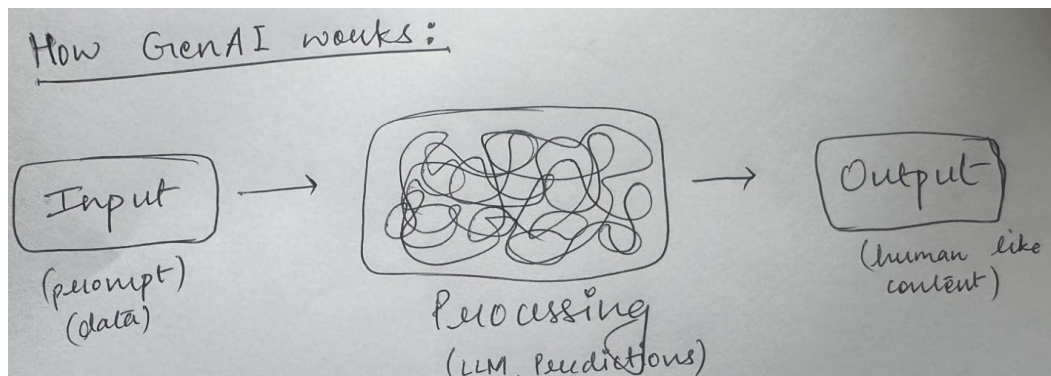


Figure 2. D1 MM illustrates how GenAI generates outputs based on a prediction model but simplistically explains complex processing.

1.1.4 Opaque technical complexity (5/9)

While five drawings represented the inputs and outputs into the system, they lacked representations of information processing within the system, suggesting that it is difficult or impossible to understand the workings of this complex system (see example in Figure 2).

The mysteriousness of the information processing within the system was also represented by the lack of details in participants' explanations of drawings:

'Despite a linear flow from input to output, the processing of the system goes through a complicated and deep process'. D4

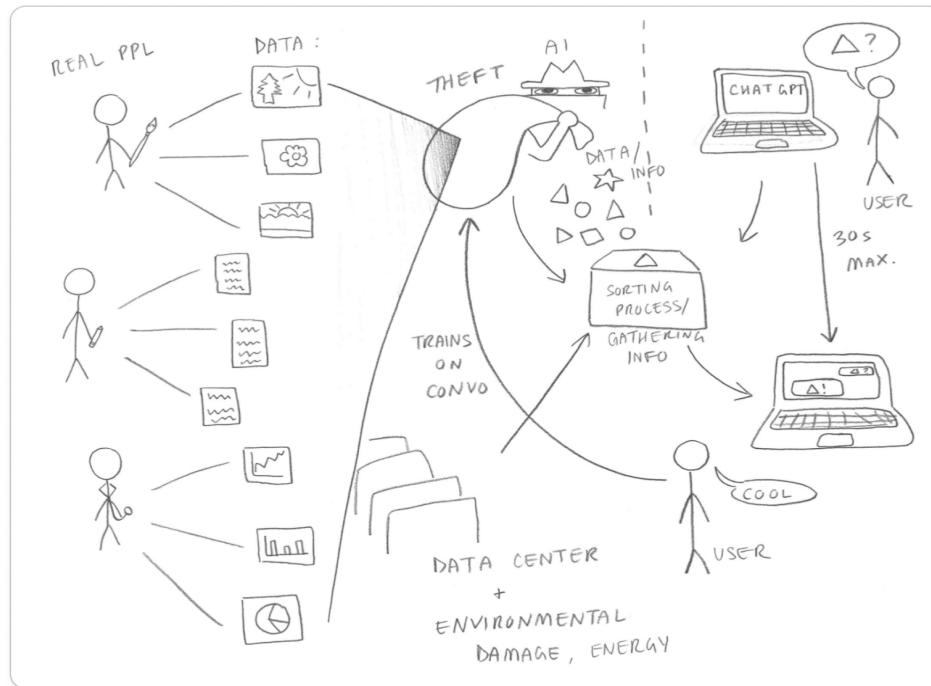


Figure 3. D8 MM represents GenAI use of information from the human-created collections as theft.

1.1.5 Presence of hardware (4/9)

Four drawings acknowledged the necessity of hardware to interface with or support GenAI (examples in Figures 1, 3, and 7):

'...then I have a drawing of, like, a computer and some servers...'. D6

1.1.6 Input and output (4/9)

Similar to the understanding of the collections and hardware that is needed to run GenAI, participants also illustrated their knowledge of clearly defined inputs, outputs, and a flow between them (Figure 1):

'The user provides a prompt... the computer highlights and encodes the prompt, then processes it through multiple networks... it comes back together... and produces a text or graphic on the screen'. D7

1.2 Misconceptions

1.2.1 Retrieval not generation (4/9)

Inaccurate representations of GenAI finding and presenting existing information, as opposed to content creation or generation, were included in four drawings (Figure 4). Participants' explanations confirmed their misconception of GenAI as a retrieval system, for example:

'[My drawing depicts] how it's trained to, maybe, retrieve from current and updated data.' D9

1.2.2 Anthropomorphism (3/9)

Three participants described GenAI as having anthropomorphic features, for example, participant D8 represented GenAI's use of existing collections in the form of a thief (Figure 3):

Another participant used anthropomorphic language in describing the system:

'...it just receives and **knows** a lot of information... then it **studies** it'. D2

This quote and drawings like Figure 3 illustrate participants' attribution of human characteristics or behaviors to GenAI.

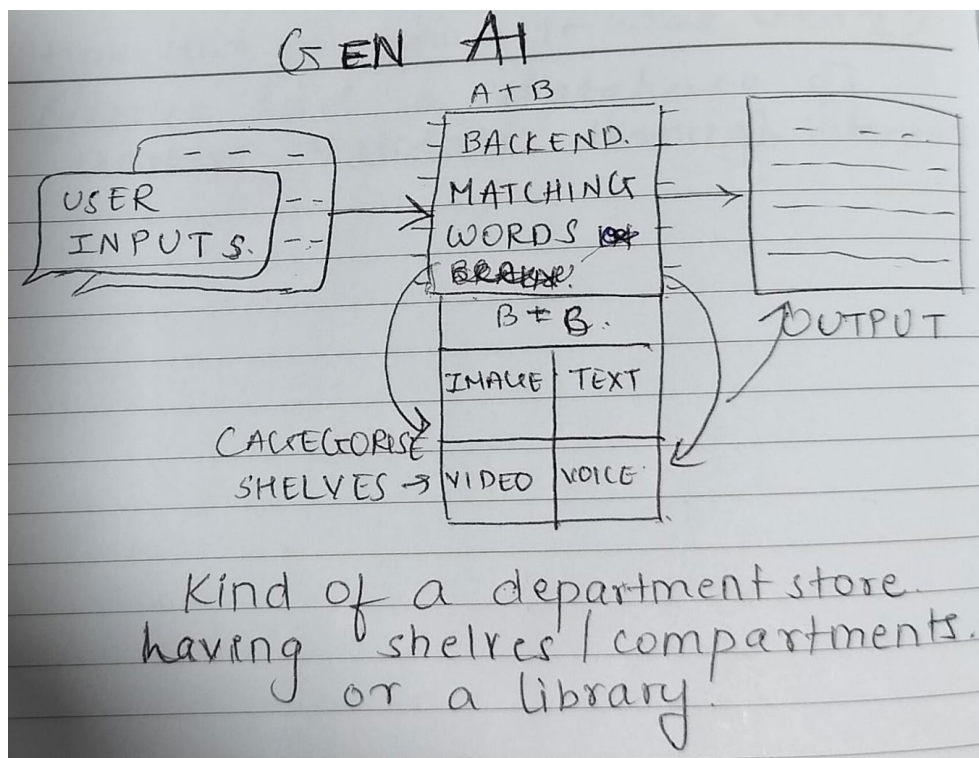


Figure 4. D4 MM describes GenAI as retrieving organised information.

Theme 2: The human

2.1 Presence of a human (5/9)

Five drawings (see examples in Figures 1, 3, and 5) included direct representations of humans, usually as end-users, in relation to the system (in their drawings, participants used the words 'user', 'client', or 'I'). One participant included humans as creators of collections used by AI (Figure 3).

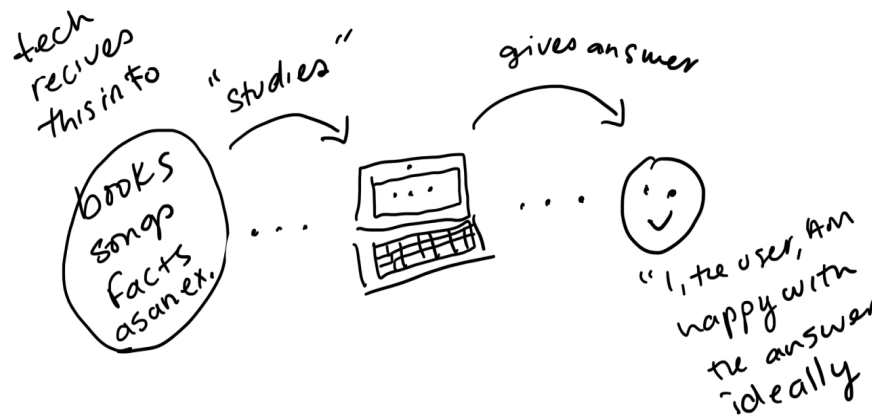


Figure 5. D2 shows a user with no interaction or influence within the system.

2.2 Absent human (4/9)

Four mental models did not contain any representations of a human in drawings or verbal descriptions (see example in Figures 2 and 6).

Theme 3: Value judgments

Three participants' mental models included references to the judgments about the function or use of GenAI.

3.1 Ethical concerns (3/9)

Ethical concerns about what is *right* or *wrong* about GenAI were depicted in three drawings (Figure 3, for example) and described by the participants:

'I drew art, writing, science, and research, and then added a little thief stealing all of it'. D8

3.2 Environmental concerns (2/9)

Two participants expressed understanding of the environmental impact of GenAI in their mental models (Figures 3 and 6):

'I also drew a data centre showing how this process consumes energy, causes heating, and leads to environmental damage'. D8

'There's a lot happening, such as how the model was trained, how it retrieves updated data, and even environmental impacts'. D9

3.3 Quality concerns (1/9)

One participant's MM (Figure 7) represented concern for the deteriorating quality of AI-contaminated training collections:

'...I used red and a little bit of it to show that it has more like, this bleed of, like, the AI into this work that's producing. ... I added red to show that AI created data becomes part of the system, which is then used in future outputs. So, it's a cycle that grows more and more'. D6

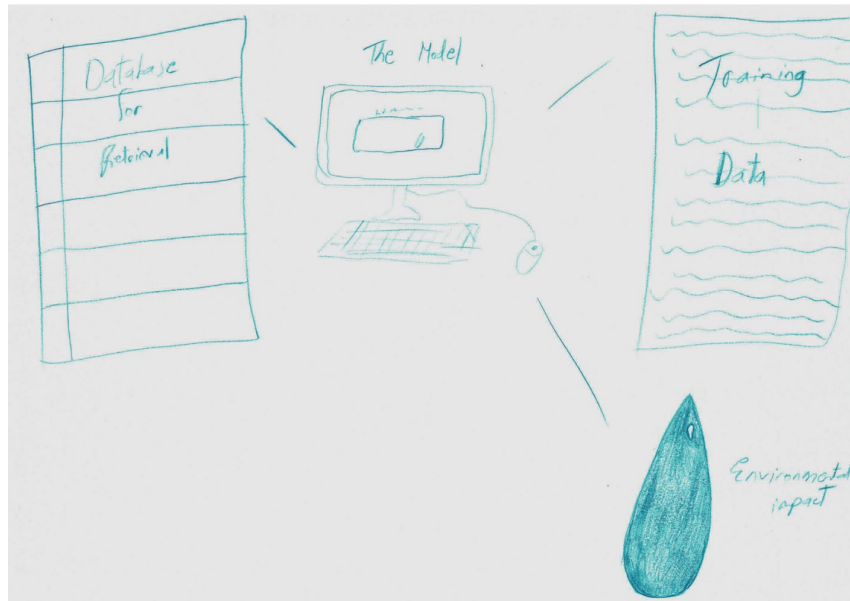


Figure 6. D9 includes a user entry point but lacks human representation; this MM also represents the environmental impact of GenAI.

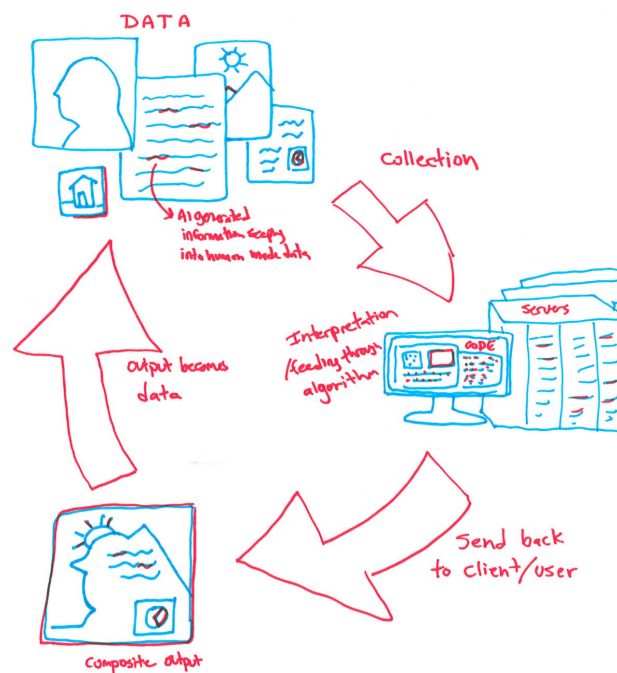


Figure 7. D8's MM represents GenAI's use of information generated from human-created collections, with the GenAI output becoming part of the collections.

Discussion

As part of the study, nine graduate students participated in two focus groups, in which they were asked to draw and discuss their mental models of GenAI. Three families of themes emerged from the qualitative data. The system family included the largest number of themes (the most grounded), which is not surprising considering that participants were asked to represent their understanding of the workings of the GenAI system. The themes that were present, or absent, from the human and value judgment families added nuance to the understanding of the students' mental models of GenAI.

Participants demonstrated a fairly accurate high-level understanding of GenAI as a pattern-seeking probabilistic system that relies on human-created (training) collections, and includes hardware, input, and output components.

Several participants incorrectly portrayed GenAI as an information retrieval system rather than a generation system, attributing human-like qualities to it.

These findings align with previous research: initial GenAI mental models could be simplified (Marx et al., 2023; Zirenko et al., 2025) and include erroneous or anthropomorphic features (Sanusi et al., 2024). Participants drew and discussed selected parts of the GenAI system, with many of the complex inner workings of the system omitted or represented abstractly (e.g., wiggly lines, connected dots). Such omissions and general representations of the system complexity point to the cognitive limitations in understanding of the opaque systems (Bansal et al., 2019), levels of technological literacy (Druga et al., 2022; Long & Magerko, 2020), and poor explainability of GenAI systems (Ehsan et al., 2022). The informal, intuitively grounded ideas represented in the mental models of the algorithmic systems can also be understood through the lens of folk theories (DeVito, 2021) and their role in constructing mental models that are grounded in experience, observation, and interaction with the platform.

The transfer of the information retrieval model to the information generation context could be explained by the human tendency to project the MM of the known systems into the unknown contexts (Norman, 1983). While assigning human features to the system is not accurate, it also represents a known trend of anthropomorphising systems to explain and predict their behaviours (Epley, Waytz, & Cacioppo, 2007). The use of anthropomorphic or familiar analogies has been shown to ease users' initial understanding of the GenAI systems (Sanusi et al., 2024). Additionally, GenAI systems are naturally human-like, with outputs mimicking the human-created data they are trained on. This may further encourage users to project familiar but potentially inaccurate mental models onto them (Bandi et al., 2023).

Human stakeholders were present in five mental models: four participants included end-users in their drawings and explanations, and one included humans as producers of the collections used by GenAI. Notably, none of the participants included humans as engineers of GenAI, suggesting that they perceive the systems as autonomous, detached from the human designers and engineers who create them. This perception might stem from limited transparency and the tendency to attribute independent agency to algorithmic systems (Nyholm, 2018; Langer et al., 2021).

Three participants assigned *value judgments* to the GenAI models, demonstrating their awareness of ethical and environmental concerns associated with GenAI technology. While these concerns among student GenAI users have been previously documented (Zhu et al., 2024; Baek et al., 2024), one of our participants expressed a deeper understanding of the adverse effects of AI-generated content 'polluting' the training collections and decreasing the overall quality of generated information. Another participant framed the issue critically, describing training data as stolen, highlighting concerns about the authorship of content.

Our exploratory study relied on a small sample of students from a single academic department. The study employed self-report methods, which, on one hand, might have obscured some of the students' opinions and behaviours, but on the other hand, it aligns with the study's interpretivist goal of capturing rich, authentic accounts from those most engaged with GenAI (Patton, 2015; Lincoln & Guba, 1985). Despite these and other limitations that restrict the generalisability of the findings, the study confirms and expands prior work on student adoption of GenAI and on mental models of information technology in general, and offers promising directions for further research.

Conclusion

As part of a larger study aimed at examining iSchool graduate students' adoption of GenAI tools, we explored students' mental models of these systems. All participants reported using GenAI systems with varying frequencies, most commonly utilising a multi-purpose ChatGPT artificial intelligence tool for academic tasks, primarily writing or brainstorming.

Asking graduate students to draw and explain their mental models facilitated informative and engaging discussions about their understanding of GenAI. Participants' familiarity with GenAI was demonstrated by the MM that included elements of AI's global behaviour (e.g., understanding GenAI as systems that generate outputs from user inputs) and knowledge distribution (e.g., recognising that GenAI models are trained on large datasets as well as user interactions, Gero et al., 2020). Some mental models accurately included elements of GenAI systems, such as probabilistic algorithms, reliance on hardware, training collections, the presence of inputs and outputs, and the general flow of information. Three participants extended their mental models to include the environmental and social impacts of GenAI. Students demonstrated a high-level understanding of GenAI complexity, although they often omitted descriptions of its detailed technical components. Students occasionally applied their knowledge of prior technology, specifically information retrieval systems, to explain the generative system. While some of the mental models included representations of end-users, creators of information collections, and occasionally, anthropomorphic system features (e.g., GenAI as a thief), human designers or engineers were not present in the representations of GenAI systems.

The findings suggest that an even more technologically literate group of information science students who frequently use AI tools for routine academic tasks has exhibited limitations in understanding GenAI, similar to other previously studied groups (Marx et al., 2023; Zirenko et al., 2025; Sanusi et al., 2024). The most notable omissions of the GenAI mental models were not technical, but rather social: students' mental models did not include human creators of AI within the system. Such *black-box* perceptions obscure the underlying human decision-making embedded in system design and functionality, and therefore need to be corrected. Students' awareness of the social and environmental impacts of GenAI should also be promoted, as only one-third of the models included these elements. Some of the MM limitations may be unavoidable and explained by psychological tendencies to draw analogies from familiar, but not always adequate, models of human and technology behaviour. Because users' experiences and the broader, societal adoption of GenAI are constantly changing, the accuracy of the mental models will be changing too and should be viewed as an ongoing process, as opposed to a one-time goal (Doyle et al., 2021; DeVito, 2021).

However, many misconceptions and omissions in the mental models, which stem from the technological and social complexity of GenAI, can be mitigated to improve students' expectations, trust, and appropriate uses of these tools (Bansal et al., 2019; Makri et al., 2006; Zhang et al., 2021). Strengthening students' metacognitive skills, such as self-awareness and monitoring abilities, can be addressed through educational resources and an AI-dedicated curriculum (Marx et al., 2023; Zirenko et al., 2025; Dahri et al., 2024; Tankelevitch et al., 2024). Developers of GenAI systems should prioritise transparency and limit cognitive overload (Davis, 1989; Dahri et al., 2024) by designing explainable features (e.g., features that give users greater control over how outputs are generated, enabling them to understand the system's limitations through interaction). Media coverage, which at the early stages of innovation tends to swing between hype and fear, obscuring AI's social and ethical dimensions (Metz & Diakopoulos, 2020), should focus on presenting AI as a sociotechnical system—shaped by data, algorithms, and humans, not autonomous entities, and communicate both its capabilities and limitations (West et al., 2019). Media can support informed public discourse by involving diverse stakeholders, offering clear explanations and examples (Long

& Magerko, 2020), and bringing attention to who builds, benefits from, or is harmed by GenAI (Crawford, 2021; Eubanks, 2018).

iSchool students are not only avid users of GenAI, whose greater understanding of this technology could lead to greater efficiency and satisfaction, but they are also future GenAI designers who can gain from an accurate and critical understanding of GenAI's technological and social components.

Acknowledgements

We are grateful to our participants for their contributions to the study. We are thankful to the reviewers for their thoughtful suggestions. The research was partially supported by the Faculty Innovation Fund award from the Pratt Institute iSchool.

About the authors

Irene Lopatovska, PhD, is a professor at the School of Information, Pratt Institute. Her research examines the relationships among humans, information, and technology (see her website <https://irenelopatovska.wordpress.com/> for more details). She can be contacted at ilopatov@pratt.edu.

Conor Mack is a master's student in Information Experience Design at Pratt Institute and a UX Researcher at Exponent. He received his BFA in Photography from Rutgers University. His research interests focus on human-centered AI and the user experience of emerging technology. He can be contacted at conoramack@gmail.com.

Ellen Connors is a MLIS candidate at Pratt Institute and a Fellow at the Museum of Modern Art Library. She received her BS in Information, with a focus on User Experience Research and Design, from the University of Michigan. Her research interests include online youth and community practices and creativity in digital environments. She can be contacted at econn170@pratt.edu.

References

- Ashktorab, Z., Jain, M., Liao, Q. V., & Weisz, J. D. (2019). Resilient chatbots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). CHI '19: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3290605.3300484>
- Amoozadeh, M., Daniels, D., Nam, D., Kumar, A., Chen, S., Hilton, M., Srinivasa Ragavan, S., & Alipour, M. A. (2024). Trust in generative AI among students: An exploratory study. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1* (pp. 67–73). SIGCSE 2024: The 55th ACM Technical Symposium on Computer Science Education. ACM. <https://doi.org/10.1145/3626252.3630842>
- Baek, C., Tate, T., & Warschauer, M. (2024). “ChatGPT seems too good to be true”: College students' use and perceptions of generative AI. *Computers and Education: Artificial Intelligence*, 7, 100294. <https://doi.org/10.1016/j.caeai.2024.100294>
- Baez, C., Buchner, J., Roos, A.-L., & Schallert-Vallaster, S. (2023). *Exploring the perception of AI in learning: Unveiling the role of student and teacher motivation and self-efficacy*. Center for Open Science. <https://doi.org/10.31219/osf.io/wqvrh>

- Banks, J. (2020). Optimus primed: Media cultivation of robot mental models and social judgments. *Frontiers in Robotics and AI*, 7. <https://doi.org/10.3389/frobt.2020.00062>
- Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019). Beyond accuracy: The role of mental models in human-AI team performance. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 7, 2–11. <https://doi.org/10.1609/hcomp.v7i1.5285>
- Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet*, 15(8), 260. <https://doi.org/10.3390/fi15080260>
- Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). Shared mental models in expert team decision making. In N. J. Castellan, Jr. (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Lawrence Erlbaum Associates.
- Carroll, J. M., & Olson, J. R. (1988). Mental models in human-computer interaction. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 45–65). Elsevier.
- Chen, A., Kim, S. S., Dharmasiri, A., Russakovsky, O., & Fan, J. E. (2025, April). Portraying Large Language Models as machines, tools, or companions affects what mental capacities humans attribute to them. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1–14).
- Chiyah Garcia, F. J., Robb, D. A., Liu, X., Laskov, A., Patron, P., & Hastie, H. (2018). Explainable autonomy: A study of explanation styles for building clear mental models. In *Proceedings of the 11th International Conference on Natural Language Generation* (pp. 99–108). Association for Computational Linguistics. <https://doi.org/10.18653/v1/w18-6511>
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Daher, W., & Hussein, A. (2024). Higher education students' perceptions of GenAI tools for learning. *Information*, 15(7), 416. <https://doi.org/10.3390/info15070416>
- Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon*, 10(8), e29317. <https://doi.org/10.1016/j.heliyon.2024.e29317>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>
- DeVito, M. A. (2021). Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–38. <https://doi.org/10.1145/3476080>
- Digital Education Council. (2024, August 2). *Digital education council global AI student survey 2024*. [Digitaleducationcouncil.com](https://www.digitaleducationcouncil.com). <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024>
- Doyle, P. R., Clark, L., & Cowan, B. R. (2021). What do we see in them? Identifying dimensions of partner models for speech interfaces using a psycholexical approach. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). CHI '21: CHI

- Conference on Human Factors in Computing Systems. ACM.
<https://doi.org/10.1145/3411764.3445206>
- Druga, S., Christoph, F. L., & Ko, A. J. (2022). Family as a third space for AI literacies: How do children and parents learn about AI together? In *CHI Conference on Human Factors in Computing Systems* (pp. 1–17). CHI '22: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3491102.3502031>
- Ehsan, U., Wintersberger, P., Liao, Q. V., Watkins, E. A., Manger, C., Daumé III, H., Riener, A., & Riedl, M. O. (2022). Human-Centered Explainable AI (HCXAI): Beyond opening the black-box of AI. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (pp. 1–7). CHI '22: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3491101.3503727>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Gero, K. I., Ashktorab, Z., Dugan, C., Pan, Q., Johnson, J., Geyer, W., Ruiz, M., Miller, S., Millen, D. R., Campbell, M., Kumaravel, S., & Zhang, W. (2020). Mental models of AI agents in a cooperative game setting. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). CHI '20: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3313831.3376316>
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, 4(4), 799–819. <https://doi.org/10.1109/tai.2022.3194503>
- Hsiao, C.-H., & Tang, K.-Y. (2024). Beyond acceptance: An empirical investigation of technological, ethical, social, and individual determinants of GenAI-supported learning in higher education. *Education and Information Technologies*, 30(8), 10725–10750. <https://doi.org/10.1007/s10639-024-13263-0>
- Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348. <https://doi.org/10.1016/j.caeai.2024.100348>
- Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20(1). <https://doi.org/10.1007/s40979-024-00149-4>
- Katsantonis, A., & Katsantonis, I. G. (2024). University students' attitudes toward artificial intelligence: An exploratory study of the cognitive, emotional, and behavioural dimensions of AI attitudes. *Education Sciences*, 14(9), 988.
- Kelly, S., Kaye, S.-A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, 77, 101925. <https://doi.org/10.1016/j.tele.2022.101925>

- Kim, J., & Cho, Y. H. (2023). My teammate is AI: Understanding students' perceptions of student-AI collaboration in drawing tasks. *Asia Pacific Journal of Education*, 45(3), 1013–1027. <https://doi.org/10.1080/02188791.2023.2286206>
- Kumar, D. R., & Raman, R. (2022). Student perceptions on artificial intelligence (AI) in higher education. In *2022 IEEE Integrated STEM Education Conference (ISEC)* (pp. 450–454). 2022 IEEE Integrated STEM Education Conference (ISEC). IEEE. <https://doi.org/10.1109/isec54952.2022.10025165>
- Langer, M., Körner, A., & Banse, R. (2021). In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human-AI interaction. *Journal of Computer-Mediated Communication*, 26(6), 384–400. <https://doi.org/10.1093/jcmc/zmab011>
- Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). CHI '20: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3313831.3376590>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
- Lopatovska, I., Mack, C., Zhu, P., Connors, E., Keller, N., Kaur, S., Koonamparampath, M., Everett, M., Zapata, Z., Giordano, P., Yang, H., Kim, M., & Shetty, N. (in press). The normalization of unease: The case study of GenAI use by iSchool students. *Information and Learning Sciences*. <https://doi.org/10.1108/ILS-08-2025-0125>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). CHI '20: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3313831.3376727>
- Long, D., & Magerko, B. (2020). What is AI literacy? *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(9), 13681–13688. <https://doi.org/10.1145/3313831.3376727>
- Maita, I., Saide, S., Putri, A. M., & Muwardi, D. (2024). Pros and cons of artificial intelligence-ChatGPT adoption in education settings: A literature review and future research agendas. *IEEE Engineering Management Review*, 52(3), 27–42. <https://doi.org/10.1109/emr.2024.3394540>
- Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C., & Buchanan, G. (2006). A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *Journal of the American Society for Information Science and Technology*, 58(3), 433–445. <https://doi.org/10.1002/asi.20510>
- Marx, E., Leonhardt, T., & Bergner, N. (2023). Secondary school students' mental models and attitudes regarding artificial intelligence - A scoping review. *Computers and Education: Artificial Intelligence*, 5, 100169. <https://doi.org/10.1016/j.caeai.2023.100169>
- Marrone, R., Zamecnik, A., Joksimovic, S., Johnson, J., & De Laat, M. (2024). Understanding student perceptions of artificial intelligence as a teammate. *Technology, Knowledge and Learning*, 30(3), 1847–1869. <https://doi.org/10.1007/s10758-024-09780-z>
- McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher

- education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 3, 100121.
<https://doi.org/10.1016/j.chbah.2025.100121>
- Metz, C., & Diakopoulos, N. (2020). The rise of AI journalism. *Columbia Journalism Review*.
<https://www.cjr.org>
- Nyholm S. (2018). Attributing agency to automated systems: Reflections on human-robot collaborations and responsibility-loci. *Science and engineering ethics*, 24(4), 1201–1219.
<https://doi.org/10.1007/s11948-017-9943-x>
- Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale, NJ: Lawrence Erlbaum
- Nourani, M., Roy, C., Block, J. E., Honeycutt, D. R., Rahman, T., Ragan, E., & Gogate, V. (2021). Anchoring bias affects mental model formation and user reliance in explainable AI systems. In *Proceedings of the 26th International Conference on Intelligent User Interfaces* (pp. 340–350). ACM. <https://doi.org/10.1145/3397481.3450639>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.
- Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Reiter, L., Jörling, M., Fuchs, C., & Böhm, R. (2025). Student (mis)use of generative AI tools for university-related tasks. *International Journal of Human-Computer Interaction*, 41(19), 12390–12403. <https://doi.org/10.1080/10447318.2025.2462083>
- Rutjes, H., Willemsen, M., & IJsselsteijn, W. (2019). Considerations on explainable AI and users' mental models. In *CHI 2019 Workshop: Where is the Human? Bridging the Gap Between AI and HCI*. ACM.
- Ronge, R., Maier, M., & Rathgeber, B. (2025). Towards a definition of generative artificial intelligence. *Philosophy & Technology*, 38(1). <https://doi.org/10.1007/s13347-025-00863-y>
- Sanusi, I. T., Ayanwale, M. A., & Tolorunleke, A. E. (2024). Investigating pre-service teachers' artificial intelligence perception from the perspective of planned behavior theory. *Computers and Education: Artificial Intelligence*, 6, 100202.
<https://doi.org/10.1016/j.caeai.2024.100202>
- Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics. *Journal of Database Management*, 31(2), 74–87. <https://doi.org/10.4018/jdm.2020040105>
- Tankelevitch, L., Kewenig, V., Simkute, A., Scott, A. E., Sarkar, A., Sellen, A., & Rintel, S. (2024). The metacognitive demands and opportunities of generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–24). CHI '24: CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3613904.3642902>
- Wang, Q., Saha, K., Gregori, E., Joyner, D., & Goel, A. (2021). Towards mutual theory of mind in human-AI interaction: How language reflects what students perceive about a virtual

- teaching assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM.
- Wang, L., & Li, W. (2024). The impact of AI usage on university students' willingness for autonomous learning. *Behavioral Sciences*, 14(10), 956.
<https://doi.org/10.3390/bs14100956>
- Wang, Q. (2024). *Mutual theory of mind for human-AI communication in AI-mediated social interaction* (Doctoral dissertation, Ph. D. Dissertation. Georgia Institute of Technology).
- West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race, and power in AI*. AI Now Institute. <https://ainowinstitute.org/discriminatingystems.html>
- Zhang, Y., Wu, M., Tian, G. Y., Zhang, G., & Lu, J. (2021). Ethics and privacy of artificial intelligence: Understandings from bibliometrics. *Knowledge-Based Systems*, 222, 106994.
<https://doi.org/10.1016/j.knosys.2021.106994>
- Zhu, W., Huang, L., Zhou, X., Li, X., Shi, G., Ying, J., & Wang, C. (2024). Could AI ethical anxiety, perceived ethical risks and ethical awareness about AI influence university students' use of generative AI products? An ethical perspective. *International Journal of Human-Computer Interaction*, 41(1), 742–764. <https://doi.org/10.1080/10447318.2024.2323277>
- Zirenko, M., Machura, I. A., Fabriz, S., Schulze-Vorberg, L., & Horz, H. (2025). AI and learning with AI: University students' metaphorical conceptualizations. *Journal of Interactive Media in Education*, 2025(1).

© [CC-BY-NC 4.0](#) The Authors. For more information, see our [Open Access Policy](#).