

Choose Your Own A(I)dvventure: Design, Development, and Evaluation of a Rule-Based Serious Game to Foster AI Risk Awareness in Students.

Pauline Klimas-Koch, Antje Neve , Dieter Pawelczak 

Department of Electrical and Computer Technology and Communication Technology, University of the Bundeswehr Munich, Germany

How to cite: Klimas-Koch, P; Neve, A.; Pawelczak, D. (2026). Choose Your Own A(I)dvventure: Design, Development, and Evaluation of a Rule-Based Serious Game to Foster AI Risk Awareness in Students. In: 12th International Conference on Higher Education Advances (HEAd'26). Valencia, 15-18, 2026. <https://doi.org/10.4995/HEAd26.2026.21747>

Abstract

The widespread adoption of generative AI in education has created an urgent need to foster students' awareness of AI-related risks beyond purely informational approaches. This paper presents the design, development, and evaluation of a rule-based narrative serious game implemented with Twine to support AI risk awareness and critical reflection. The game models everyday AI-related decision-making and deliberately contrasts the perceived convenience of AI use with its potential consequences, such as hallucinations and cognitive offloading. Following a design science research methodology, the prototype was evaluated in an exploratory quasi-experimental pilot study with upper secondary students and compared to a dialog-based RAG chatbot. Results suggest that while both approaches support short-term factual learning, the rule-based game shows indications of fostering a more differentiated and critical attitude toward AI use, whereas the dialog-based system primarily increases perceived safety and confidence.

Keywords: Artificial Intelligence, Serious Games, AI literacy, Twine, Narrative Games.

1. Introduction

The rapid diffusion of generative Artificial Intelligence (AI) has fundamentally altered learning practices in secondary and higher education. Empirical data illustrate the exceptional speed with which AI tools have transitioned from niche technologies to everyday educational instruments. In Germany, approximately 62% of adolescents aged 12–19 reported using AI applications in 2024; by 2025, this share had risen to 91% (Feierabend et al., 2025). A comparable trend is observed among university students. Despite this widespread adoption, learners aged 14–27 often lack a systematic understanding of the risks associated with generative AI. This

ambivalence arises from the fact that users benefit from emergent capabilities, while simultaneously being exposed to risks inherent in probabilistic token prediction, including hallucinations, bias and data misuse. When cognitive processes are uncritically delegated to AI systems - a phenomenon described as *cognitive offloading* - learning outcomes may be undermined in the long term. Addressing this tension constitutes a central didactic challenge for contemporary education (OECD, 2026). This work shows how a *Serious Game* implementation can enhance the students' ability to critically use AI systems.

2. Related Work – Cognitive Debt and Serious Games

Recent exploratory studies suggest that prolonged reliance on Large Language Models (LLMs) may negatively affect recall and cognitive engagement (Kosmyna et al., 2025). While the concept of *cognitive debt* is still emerging, it highlights potential long-term risks of uncritical AI use. Institutional reports echo these concerns. The OECD (2026) warns that excessive reliance on AI tools may increase task performance while simultaneously weakening deep learning processes. To counteract this development, AI should not be used as a shortcut but embedded in intentionally designed learning scenarios. However, many existing educational interventions remain predominantly informational, focusing on guidelines, ethical principles, or technical explanations. Such approaches struggle to translate abstract risks into personally meaningful experiences, particularly for digitally fluent learners accustomed to seamless AI assistance.

Serious Games enable learners to experience the implications of their choices within a protected digital space. Abstract risks such as algorithmic bias or data leakage can thus be transformed into concrete narrative events (Jenkins, 2014). Unlike many existing educational games that focus on isolated subtopics, this work addresses the everyday use of AI tools in typical situations. The objective is not to discourage AI usage, but to cultivate a reflective and situationally appropriate form of engagement. The open-source tool Twine enables educators and learners to collaboratively develop and adapt narrative scenarios without programming expertise, supporting participatory and sustainable curriculum development (Harvey, 2014). Twine provides fully deterministic learning paths: all scenarios, consequences, and feedback are intentionally designed and reproducible. This allows AI-related risks to be modeled in a controlled manner without introducing additional uncertainties through the learning tool itself. Taking advantage of the mentioned features the game „Choose Your Own A(I)dventure“ was implemented using Twine. Due to its web-based concept, it can easily be facilitated into school and university contexts.

3. Research Question and Methodological Design

To evaluate the learning effect of the rule-based Twine game, a dialog-based RAG chatbot is developed and used for the control group. This study follows the Design Science Research Methodology (Peppers et al., 2007). Problem identification addresses the limited availability of experiential approaches to AI risk education, from which educational objectives focusing on risk awareness and reflective AI use were derived. The effects of the Twine Game and the Chatbot were analyzed in an exploratory quasi-experimental pre-post study, followed by a critical discussion of limitations and design implications. The conceptual starting point is the technological ambivalence of LLMs, which combine high practical utility with inherent risks. From this ambivalence, two learning objectives were derived: strengthening AI risk awareness (e.g., recognizing hallucinations or bias in concrete situations) and supporting reflective, context-sensitive decision-making when using AI tools under real-world constraints. The goal is to support neither uncritical acceptance nor complete avoidance of AI-Tools, but to support a reflective, context-sensitive use of AI while avoiding a didactic or admonishing tone. The game creates experiential spaces in which consequences resulting from uncritical AI use can be simulated without real-world repercussions.

4. Pedagogical and Narrative Design

The pedagogical design of the game is specifically aligned with Kolb's (1984) experiential learning model. Learning is initiated through concrete decision-making situations embedded in realistic everyday scenarios. Consequence-based feedback supports reflection on causal relationships between AI use and its potential risks. Rather than progressing linearly through the learning cycle, the game enables iterative exploration: players are encouraged to replay scenarios and experiment with alternative choices, allowing insights to be consolidated through comparison and contrast. From a narrative perspective, the game follows a blending paradigm (Hoblitz, 2015) by embedding educational objectives implicitly within the gameplay. To enhance immersion and personal involvement, scenarios are presented in second-person narration (Salter, 2016). Such narrative involvement facilitates identification and reflective engagement - a correlation well-established in the field of interactive storytelling (Jenkins, 2014). The game consists of five independent but thematically connected everyday scenarios, with each scenario representing one specific pre-defined AI risk dimension (see Table 1).

Table 1. Scenarios and associated AI risks.

Scenario	AI Risk Dimension	Key Source
The Homework	Hallucinations	Klimova & Pikhart (2025)
The Coding Class	Cognitive Offloading	Kosmyna et al. (2025)

Choose Your Own A(I)dventure

The Classmate	Bias	Bundesamt für Sicherheit in der Informationstechnik [BSI] (2025)
The Job Application	Data Leakage / Privacy	Staab et al. (2024)
The Band Contest	Plagiarism	Hutson (2024)

In each scenario, players have access to a fictional AI assistant (see Figure 1a) and must decide whether and how to use it. AI use is intentionally non-binary: selecting the tool typically leads to further decisions regarding verification, modification, or uncritical adoption of AI-generated output. Choices made within a scenario affect its outcome and inform the player's overall competence profile, but do not prevent exploration of alternative paths through replay (see Figure 1b). To support reflection, the game provides contextual feedback and optional background information on AI-related risks such as hallucinations, bias, or cognitive offloading.

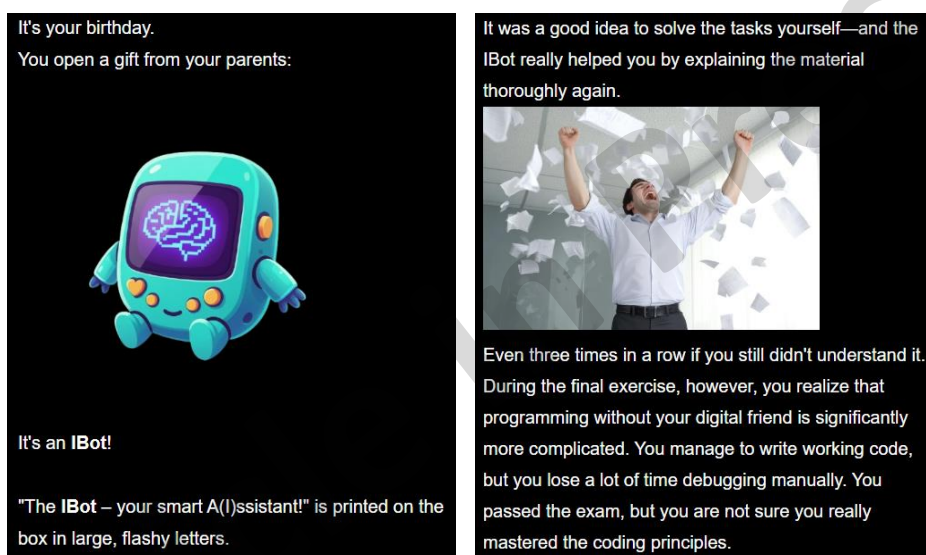


Figure 1a. Screenshot from Starting Page and 1b Scenario 2 “The Coding Class” End 3 of the Twine game.

These materials are integrated as in-game resources rather than external explanations, reinforcing the experiential character of the learning process. At the end of the game, players receive structured feedback in the form of a competence profile based on predefined risk dimensions. The underlying classification logic is fully deterministic and rule-based, ensuring transparent and reproducible feedback (see Figure 2).

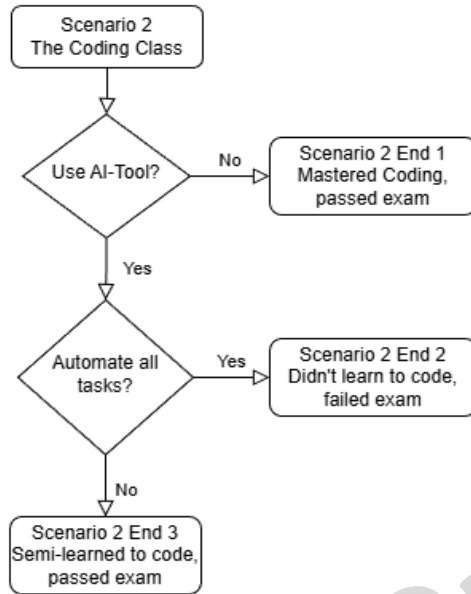
Choose Your Own A(I)dventure

Figure 2. Scenario 2 “The Coding Class” flow chart from the Twine game.

The standalone RAG-based chatbot draws on the same curated knowledge base as the Twine game and guides learners through narrative scenarios in a dialog-based format, assigning a competence profile at the end of the interaction.

5. Empirical Evaluation

The effectiveness of the two educational artifacts was explored in an exploratory quasi-experimental pilot study with 19 upper secondary students in two randomly assigned groups. Group A interacted with the rule-based Twine game ($n = 9$), while Group B used the dialog-based RAG chatbot ($n = 10$). A pre–post design was employed to capture short-term changes in knowledge and attitudes toward AI. Due to the small sample size, the analysis is descriptive and does not involve inferential statistical testing. Knowledge-related outcomes were measured using multiple-choice items addressing key AI risk dimensions. Both groups showed learning gains in the post-test, indicating short-term acquisition of factual knowledge.

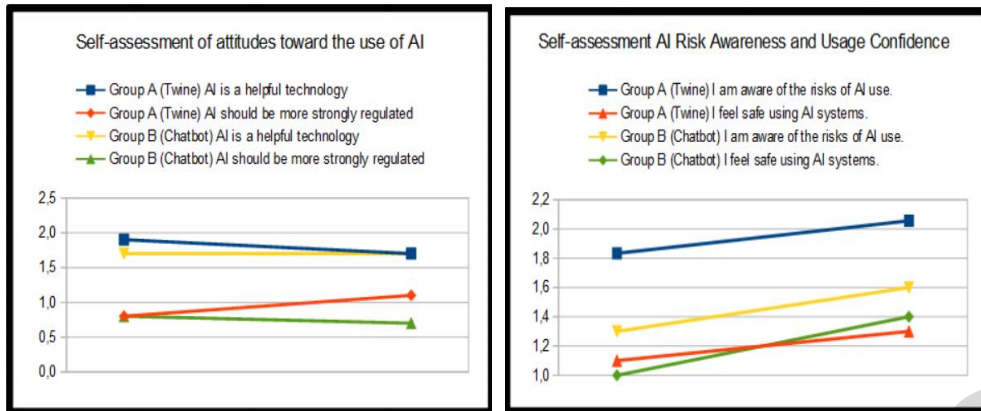
Choose Your Own A(I)dventure

Figure 3a & b. Self-assessment attitude towards AI and risk awareness.

Attitudinal changes were assessed using self-report items on a five-point Likert scale (fully agree 2.5 to –2.5 fully disagree). Participants in the Twine condition showed tendencies toward increased support for AI regulation and higher self-assessed reflectiveness in AI use, alongside a slight decrease in agreement with statements emphasizing the general helpfulness of AI. In contrast, participants in the chatbot condition primarily reported increased perceived safety and confidence when using AI tools, while showing little change in reflective self-assessment. (cmp. Fig. 3a & b). Taken together, these patterns suggest that the two approaches support different aspects of AI literacy. While both appear suitable for short-term knowledge acquisition, the rule-based game shows exploratory indications of fostering reflective engagement with AI-related risks. The dialog-based chatbot, by contrast, seems to strengthen perceived competence and safety, potentially without a corresponding increase in critical self-reflection. These observations should be interpreted cautiously given the exploratory nature of the study and the reliance on self-reported measures.

5.1. Discussion

The findings highlight how technological implementation shapes learning outcomes in AI education. In line with the research question, the deterministic structure of the Twine game appears to support reflection by making the consequences of AI-related decisions explicit and reproducible. This aligns with pedagogical arguments emphasizing the importance of didactic control when addressing complex and potentially misleading technologies. Conversely, the open-ended interaction afforded by the generative chatbot may blur the boundary between assistance and authority, reinforcing users' confidence without necessarily encouraging critical distance. While such systems can efficiently communicate information and simulate conversational engagement, they may be less suitable as standalone tools for fostering reflective judgment about AI risks. From a design perspective, the results support the view that different

educational objectives require different technological affordances. Rule-based narrative environments can be particularly effective when the goal extends beyond factual understanding toward reflective and responsible AI usage.

5.2. Limitations

This study is subject to several limitations. The small sample size and quasi-experimental design preclude statistical generalization and limit the findings to exploratory tendencies. Moreover, the intervention captured only immediate effects following a single interaction; long-term retention and stability of attitudinal changes remain unexamined. Finally, several measures rely on self-reported data, which may not fully reflect actual behavioral competence in real-world AI use.

6. Conclusion and Outlook: Toward Hybrid Learning Systems

The study reveals differentiated effectiveness profiles for distinct approaches to AI education. While both prototypes successfully conveyed factual knowledge, the rule-based Twine game showed exploratory indications of supporting a more critical stance. Its deterministic warning scenarios functioned as a didactic safety anchor, sensitizing learners to regulatory needs and the limits of AI reliability. In contrast, free interaction with the chatbot tended to increase perceived safety and technological confidence without a comparable rise in critical awareness. Given that neither approach fully addresses all dimensions of AI literacy, a hybrid didactic architecture appears promising. An initial exploratory phase using dialog-based systems may efficiently introduce concepts, followed by a structured confrontation phase in a rule-based environment where learners must apply knowledge and experience the consequences of flawed AI output. In this way, the technical limitations of generative systems become the subject of learning itself, supporting the development of stable, critically reflective AI competence.

References

- Bundesamt für Sicherheit in der Informationstechnik. (2025). *Bias in der künstlichen Intelligenz* [Bias in artificial intelligence]. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Whitepaper_Bias_KI.pdf?__blob=publicationFile&v=5
- Feierabend, S., Rathgeb, T., Gerigk, Y., & Glöckler, S. (2025). *JIM-Studie 2025: Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. [JIM Study 2025: Youth, Information, Media. Basic survey on the media usage of 12- to 19-year-olds.] Medienpädagogischer Forschungsverbund Südwest. <https://www.mpfs.de>
- Harvey, A. (2014). Twine's revolution: Democratization of game design. *Loading... The Journal of the Canadian Game Studies Association*, 8(13), 92–99. <https://journals.sfu.ca/loading/index.php/loading/article/view/155>

- Hoblitz, A. (2015). *Spielend lernen im Flow: Die motivationale Wirkung von Serious Games im Schulunterricht*. [Learning playfully in Flow: The motivational impact of serious games in school classrooms.] Springer VS. <https://doi.org/10.1007/978-3-658-11376-6>
- Hutson, J. (2024). Rethinking plagiarism in the era of generative AI. *Journal of Intelligent Communication*, 4(1). <https://doi.org/10.54963/jic.v4i1.220>
- Jenkins, K. (2014). *Choose your own adventure: Interactive narratives and attitude change* [Master's thesis, University of North Carolina at Chapel Hill]. Carolina Digital Repository. <https://doi.org/10.17615/6MA7-4348>
- Klimova, B., & Pikhart, M. (2025). Exploring the effects of artificial intelligence on student and academic well-being in higher education: A mini-review. *Frontiers in Psychology*, 16, Article 1498132. <https://doi.org/10.3389/fpsyg.2025.1498132>
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks* (arXiv Preprint). <https://arxiv.org/abs/2506.08872>
- OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Salter, A. (2016). Playing at empathy: Representing and experiencing emotional growth through Twine games. In *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)* (S. 1–8). IEEE. <https://doi.org/10.1109/SeGAH.2016.7586272>
- Staab, R., Vero, M., Balunović, M., & Vechev, M. (2024). Beyond memorization: Violating privacy via inference with large language models. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/2310.07298>