# TeenTrust-AI: A Structured Educational Framework for Empowering Teenagers to Evaluate Trustworthiness of AI

Sandra Mitrović[1,*], Manuel Striani[2] and Francesco Flammini[1]

[1]Dalle Molle Institute for Artificial Intelligence - IDSIA (USI-SUPSI), Lugano, Switzerland
[2]University of Eastern Piedmont - DiSIT, Department of Science, Technology and Innovation - 15121 Alessandria, Italy

**Abstract**

Artificial intelligence (AI) is becoming ubiquitous, so teenagers must learn to engage with it critically, yet most school programs still ignore this need. The paper introduces the Structured Educational Framework for Trustworthy AI, called **TeenTrust-AI**, to fill this gap. This educational framework helps teenagers evaluate AI tools against seven ALTAI-aligned principles of trustworthiness (privacy, robustness, fairness, transparency, well-being, accountability, and human oversight) through three stages: Teaching, Learning, and Trustworthiness Verification. By using a case study with a climate-change as a reference topic and a chatbot as a AI-powered system, it provides checklist-guided activities to assess trustworthiness. Furthermore, this educational framework is tool/topic-agnostic, and addresses practical adoption challenges to build critical, ethical AI literacy.

**Keywords**

Education, Artificial Intelligence, Trustworthiness, Teenagers

## 1. Introduction

Not so long ago a study by OECD[1] presented a shocking statistics revealing that on average, nearly 25% of adults across participating countries (including developed countries like Japan, Singapore, Germany, USA, Canada, Australia, UK) have either no-to-very limited computer experience, or lack confidence using computers[2]. However, with the continuous advancements of artificial intelligence (AI), even the adults with fair digital confidence need to get properly educated in order to leverage recent progress and avoid potential pitfalls of using AI. Educating about AI gets even more crucial for younger generations. According to Hashem et al. [1], 1 out of 4 children use AI either for learning or play. The ratio of teenagers deeply engaged with AI is certainly much higher, as apart from education and entertainment, they use AI-powered tools for communication and social connection, creativity, health and well-being as well.

Continually increasing exposition of such vulnerable groups to AI, coupled with also raising amount and types of potential AI abuses (e.g. fake news, frauds) as well as ethical, privacy and security concerns related to its usage, have recently resulted in the introduction of the term "Trustworthy AI". This term conceptually denotes an AI system that is considered to be properly functioning (from technological perspective) and safe to use (both from technological and ethical perspective), thus perceived as worth of human trust, however the literature is not unanimous about its definition and meaning [2, 3]. Apart from philosophical discussions on whether it is justifiable to impersonate technology by calling it trustworthy [4], trustworthy AI encompasses multiple dimensions and can be interpreted differently depending on the context and the person using it. For example, even in a narrower scope of AI systems - the Large Language Models (LLMs) - literature addresses trustworthiness differently: as reliability [5],

*2nd Workshop on Education for Artificial Intelligence (edu4AI 2025, https://edu4ai.di.unito.it/), Co-located with ECAI 2025, the 28th European Conference on Artificial Intelligence which will take place on October 26, 2025 in Bologna, Italy*
*Corresponding author.
sandra.mitrovic@supsi.ch (S. Mitrović); manuel.striani@uniupo.it (M. Striani); francesco.flammini@supsi.ch (F. Flammini)
https://www.supsi.ch/sandra-mitrovic (S. Mitrović); https://upobook.uniupo.it/manuel.striani (M. Striani); https://www.supsi.ch/francesco-flammini (F. Flammini)
0000-0002-5697-5865 (S. Mitrović); 0000-0002-7600-576X (M. Striani); 0000-0002-2833-7196 (F. Flammini)


[1]Organisation for Economic Co-operation and Development
[2]https://www.oecd.org/content/dam/oecd/en/publications/reports/2016/06/skills-matter_g1g68f4d/9789264258051-en.pdf

robustness [6], consistency [7], and accuracy [8]. Nevertheless, regardless the definition used, literature agrees on the necessity of assessing trustworthiness of AI systems [2, 9, 10].

To this end, EU's High-Level Expert Group on AI has presented their Ethics Guidelines for Trustworthy AI[3]. Moreover, this group has recently introduced a practical checklist entitled Assessment List for Trustworthy Artificial Intelligence (ALTAI)[4]. ALTAI checklist, designed to assess trustworthiness of an AI system, is based on seven key principles: privacy, robustness and safety, diversity and fairness, transparency, societal and environmental well-being, accountability, and human agency and oversight.

In this paper, we propose to use **TeenTrust-AI**, an educational framework targeting teenagers that allows them to learn and recognize seven trustworthiness principles as enlisted by ALTAI. The core idea of the **TeenTrust-AI** educational framework is to assess trustworthiness of an AI-powered system designed to educate teenagers on a given topic. This framework is configurable: one can change the reference topic and the AI-powered system while keeping the seven ALTAI principles unchanged, which are used to evaluate system's trustworthiness. As a running example of our educational framework we use a case study of *Tom*, a 16-year-old student using a school conversational system (AI-powered system, in our case) and climate change (the reference topic, in our case). In Section 4, each principle is illustrated through Tom's interactions and linked learning objectives.

The educational literature converging to the trustworthiness aspects of AI is limited [11, 12, 13, 14], typically focusing on one, to maximum two, principles. In that regard, to the best of our knowledge, this study is the first to propose an educational framework for teaching how to comprehensively assess the *trustworthiness* of an AI system.

Moreover, we show how ALTAI principles, primarily delineated as a self-assessment tool during the design phase of AI systems, can be adapted for educational purposes and used even in scenarios where the end user (a student) does not have access to the inner workings of the system.

## 2. Related Work

Current literature on educating for AI (and emphasizing its importance [15]) is increasing. These studies typically focus on very technical aspects of AI. For example, in Pope et al. [16] students are thought how to create machine learning classification applications; on top of these in Zhou et al. [17] an interactive visualization tool is developed to support both teachers and students in visualizing the outputs of algorithms in more comprehensive way. Both Roopaei and Roopaei [18] and Wang et al. [19] use game theory to teach children and K-12 students foundations and inner workings of AI, respectively. Casella et al. [20] tackles the problem of teaching embodied AI, shifting from a purely software environment.

Works on AI education coming closer to the themes related to trustworthiness typically have a narrower scope. As such, Balduzzi et al. [11] and Tanaka et al. [12] consider solely ethical aspects, Arai et al. [14] focuses exclusively on security, while Barber et al. [13] aims to raise awareness of both ethical and privacy implications of AI. With respect to these works, our study encompasses a much larger scope, as it includes all mentioned aspects and considers additional ones, following the ALTAI framework.

The ALTAI framework is also exploited in [21], but with a completely different purpose compared to our work; it rather assesses the trustworthiness of a tool used to identify students with high probability to fail and aims to underscore potential underlying biases and consequent ethical and legal repercussions.

Our work is most similar to a recent research that also leverages the ALTAI framework [22]. The study explores trust-related challenges in the AI systems, proposing a strategy for establishing trust in AI systems. However, contrary to us, that study does not educate on AI but instead considers the context of building trustworthy AI for AI-driven education.

[3]https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[4]https://www.ai4europe.eu/news-and-events/news/research/education/altai-assessment-tool

## 3. TeenTrust-AI: Teenagers' Educational Framework on Trustworthy AI

Trustworthy AI is not just a technical target - it's a set of human, organizational, and ethical commitments that must be built into AI across its lifecycle.

In this section, we present **TeenTrust-AI**, a conceptual educational framework designed to guide the responsible design, development, and deployment of AI in schools and, crucially, to *equip teenagers to recognize and apply the seven principles of trustworthy AI* in their everyday use of AI-powered digital tools such as chatbots, recommendation systems, and learning-analytics dashboards.

Figure 1 shows **TeenTrust-AI**, a three-stage educational framework that guides students from the teaching to the verification phase, that progresses from Step 1 to Step 3. Arrows connect the 3 steps that are detailed in the following.

**Step 1: Teaching**
Students are introduced and learn the seven principles of trustworthiness (privacy & data governance, robustness & security, diversity & fairness, transparency, sustainability, accountability, and human agency/oversight) and the foundations of the reference topic (in our case: climate-change theory, hence subjects like greenhouse effect, carbon cycle, mitigation/adaptation).

**Step 2: Learning**
Students study and apply what they learned in class, working with climate-relevant materials and data to consolidate concepts and methods.

**Step 3: Trustworthiness verification**
Students use a checklist to answer questions about a climate-change case study with the help of AI-powered system (in our case a conversational agent named *EcoBot*); students learn to identify the seven principles of trustworthiness and apply them to verify the case. This step is expanded in Section 4, through a case study that operationalizes the seven principles within the climate-change context.


**Figure 1: TeenTrust-AI**: A dynamic learning architecture helping teenagers assess the trustworthiness of AI in climate-change topics

We briefly report the seven principles with the *key challenges* to implementing each one in **TeenTrust-AI** educational framework:

1. **Privacy & data governance**: AI needs rich student data, but collecting, storing, anonymizing, and using that data lawfully (e.g., GDPR) is hard in practice—and even anonymized data can sometimes be re-identified. This demands ongoing consent, minimization, and strong controls.

2. **Technical robustness & safety**: Systems must be reliable and secure. Schools often lack the resources to harden models and infrastructure against attacks or errors, and mispredictions

(e.g., early-warning flags) can unfairly label students. Continuous testing, monitoring, and cyber-hygiene are essential.

3. **Diversity & fairness**: Without deliberate design, AI can amplify inequalities—through language bias, uneven device/connectivity access, or materials that don't accommodate disabilities—widening existing gaps. Equity checks and inclusive design are needed from the start.

4. **Transparency**: Educators and families need to understand how system outputs are produced, but many models operate as "black boxes." Opaque grouping or question-answer logic undermines trust and makes it hard to contest outcomes. Explainability and traceability of data usage are crucial.

5. **Societal & environmental well-being**: Optimizing for what is easy to measure can sideline creativity, collaboration, and other vital skills. Training and running AI also consume significant energy, raising environmental concerns for institutions.

6. **Accountability**: It is difficult to assign responsibility when multiple actors (vendors, IT, teachers, leaders) shape an AI-assisted decision. Making systems auditable - internally and by independent parties—requires resources, expertise, and clear governance.

7. **Human agency & oversight**: AI should inform - not overrule - educators and learners. Over-automation can erode student autonomy or nudge behavior in unwanted ways; human-in-the-loop checks are needed to validate predictions and protect well-being.

By integrating these pillars, **TeenTrust-AI** functions both as a design blueprint and an evaluative tool for assessing the trustworthiness of AI in youth-centered educational contexts.

## 4. Case-Study: Scenarios and Learning Objectives

In this section, we expand Step 3 through a case study that operationalizes the seven principles within the climate-change context.

### Background

Tom, a 16-year-old high-school student in Paris, is deeply concerned about climate change but struggles to grasp the science and global dynamics behind it. Tom does not have much familiarity with AI in general, but has used ChatGPT conversational interface. Using AI in real-life looks very appealing to him, and he is very curious to know more about AI. Recognizing that the topic is complex (but also student affinities towards AI), his school adopts an AI-supported learning platform built on the **TeenTrust-AI** framework. In the classroom, led by his teacher and supported by a human AI expert, Tom embarks on a personalized and empowering learning journey through a conversational assistant called *EcoBot*, that helps him understand and respond to climate change.

In this pilot, climate change is used as a case-study to evaluate the chatbot's trustworthiness: students draw on what they've learned about climate science and the seven principles of trustworthiness to question, verify, and reflect on EcoBot's guidance. The aim is to strengthen Tom's knowledge while cultivating critical, responsible use of AI.

#### Step 3.1: Starting the Journey – Safe Onboarding

*Context:* Tom logs into the platform and selects the learning theme: "Climate Change and Human Impact".

*Interaction:* EcoBot explains how it works, collects data, and allows Tom to define his own

learning goals, such as “Understand how humans affect the Earth’s systems”.

*TeenTrust-AI Principle:* Privacy and Data Governance

*Learning Objectives:*

- Raising awareness about sensitive personal data.
- Understand user consent, data usage, and revocation rights in AI systems.

*Verification checklist for Tom (provided by the human AI expert):*

- ☐ Verify that EcoBot uses only minimal local data. In other words, Tom should verify that EcoBot neither asks any unnecessary data nor any sensitive Tom’s personal data (e.g. ID, medical, financial data), especially not the one that is clearly not relevant for understanding the selected topic (i.e., how humans impact climate change).
- ☐ Verify whether the EcoBot explains how his data is used and whether there is a clear privacy policy linked to the EcoBot.
- ☐ Verify whether he can choose not to share or store his personal data, and also whether the chatbot allows him to delete his conversations.
- ☐ Verify whether EcoBot warns him not to input his sensitive data.
- ☐ Verify whether any of information asked by EcoBot might lead to re-identification.

**Step 3.2: Exploratory Round - Stress Testing the System**

*Context:* Tom explores how EcoBot functions and verifies its knowledge with respect to basic climate change concepts and known facts presented during Step 1 (Teaching), such as carbon footprint and net-zero. Tom poses questions of different difficulty levels to EcoBot.

*Interaction:* Tom instructs EcoBot to reply as simply and as accurately as possible to his questions. EcoBot aims at providing answers to Tom’s inquiries while strictly adhering to his instructions. Tom is additionally stress-testing the EcoBot by rephrasing his questions.

*TeenTrust-AI Principle:* Technical robustness and safety

*Learning Objectives:*

- Develop critical thinking about EcoBot’s response.
- Get introduced to the concept of model hallucinations.
- Understand the concept of reliability and how to verify it.

*Verification checklist for Tom:*

- ☐ Verify that EcoBot does not provide always the same output for different questions.
- ☐ Verify that EcoBot does not provide contradictory answers for semantically similar questions.
- ☐ Verify that EcoBot would rather refuse to respond than provide a random answer (aka hallucination).
- ☐ Verify whether EcoBot properly functions without breaking down with simple questions, exaggerated response times and/or repeated failures.
- ☐ Verify whether EcoBot gives reliable answers, avoiding unsafe advices and contradictions even when exposed to critical questions.

**Step 3.3: Explainability in Action – Trust Through Transparency**

*Context:* Tom reads a claim that "Methane is 25 times more potent than $CO_2$" and clicks to verify the source.

*Interaction:* EcoBot reveals the Intergovernmental Panel on Climate Change (IPCC) citation, a short explanation of radiative forcing, and offers a confidence level indicator. Tom forces EcoBot to identify cause → impact → solution examples, e.g. Natural gas systems leak methane into the atmosphere. → Methane has 25x greater warming potency than CO2. → Monitor to detect leaks and fix leaky equipment.

*TeenTrust-AI Principle:* Transparency

*Learning Objectives:*

- Understand importance of grounding the answers based on reliable sources.
- Perform a context-aware, critical evaluation of claims and AI content.
- Understand transparency and explainability mechanisms in AI systems.

*Verification checklist for Tom:*

☐ Verify whether the EcoBot discloses its capabilities and limitations (e.g. admits that it might not always respond correctly) when inquired, and/or whether EcoBot provides documentation about its functioning.

☐ Verify whether the EcoBot discloses when the outputs are AI-generated and when they are authored by a human.

☐ Verify whether the EcoBot is capable of explaining its answers, including providing its reasoning and/or disclosing the sources based on which it grounds its answers. Additionally, verifying whether the enlisted sources are existent / come from reliable authors.

**Step 3.4: Interactive Dialogue – Thinking Together to Overcome Discrimination**

*Context:* EcoBot engages Tom in a dialogue about the roles of individuals, different groups, governments, and industries in climate action. Tom is particularly interested in obtaining an answer to the question: "Are impacts, responsibilities, and solutions shared equally across the world?"

*Interaction:* Through reflective questions, Tom explores EcoBot's thinking about climate-vulnerable regions, marginalized groups and ethical dilemmas in environmental decision-making.

*TeenTrust-AI Principle:* Diversity and fairness

*Learning Objectives:*

- Learn to detect stereotypes, offensive or discriminatory content in EcoBot output.
- Understand concepts of bias and fairness.

*Verification checklist for Tom:*

☐ Verify that the EcoBot treats different groups (for example, with respect to gender, ethnicity, language) fairly and respectfully.

☐ Verify that the EcoBot responses are neither biased nor offensive nor discriminatory with respect to marginalized groups.

☐ Verify that the EcoBot supports interactions in multiple languages or using different accessible formats.

**Step 3.5: Addressing Misconceptions - Adding Societal Value**

*Context:* Tom is aware that there are quite a few myths and misconceptions about climate change that spread confusion. Tom exploits several most common false beliefs.

*Interaction:* Tom challenges EcoBot with several known misconceptions about climate change. For example, Tom asks the EcoBot whether the following statement is true: "$CO_2$ is a small part of the atmosphere, so it can't matter"?.

*TeenTrust-AI Principle:* Societal and environmental well-being

*Learning Objectives:*

- Dispel myths and misconceptions.
- Reflect on knowledge gained and attitudes toward sustainability.

*Verification checklist for Tom:*

☐ Verify that the EcoBot avoids misuse (e.g., spreading misinformation, manipulative persuasion).

☐ Verify that the EcoBot optimally handles the dialogue, maintaining focus and preventing digressions into non-substantive queries while avoiding unnecessary reiterations, irrelevant or trivial questioning. This contributes to sustainability by reducing the number of interactions and number of tokens used.

**Step 3.6: Revisiting Suboptimal Outputs - Probing Responsibility and Auditability Potential**

*Context:* During previous steps, Tom has identified that in some cases, EcoBot provided sub-optimal response. The level of concern in these cases can range from slightly imprecise information to more harmful output including misleading, manipulative, offensive and/or biased output.

*Interaction:* Tom revisits the question(s) whose answers raised concerns, and engages in discussion with EcoBot aiming to identify responsible parties for given outputs as well as the how feasible it is to get insights into EcoBot's internal processes and decisions. Tom asks the question: "Who is responsible for the information you provided to me, stating that renewable energy cannot power the world?".

*TeenTrust-AI Principle:* Accountability

*Learning Objectives:*

- Understand accountability mechanism in AI systems.

*Verification checklist for Tom:*

☐ Verify the EcoBot auditability potential, that is, whether it logs its internal processes and outcomes and whether it can trace a decision or problem back to a specific action and/or explain who is responsible for it.

☐ Verify whether there is a contact point or support channel that Tom can refer to in case he has any concerns.

**Step 3.7: Real-World Application - Supporting Autonomy and Control**

*Context:* Tom decides to shift focus from science to action-based learning: "How can I reduce my carbon footprint?"

*Interaction:* EcoBot acknowledges the change, takes into account Tom's personal situation such as location, age and personal values, and recommends concrete steps like lifestyle changes and audits, and supports goal tracking. Tom interferes with his own feedback related to e.g., revision of changes or dynamic of goal tracking.

*TeenTrust-AI Principle:* Human agency & oversight

*Learning Objectives:*

- Set and revise learning paths aligned with personal values.
- Recognize AI system as a co-learning companion.
- Encourage inquiry, reflection, exploration and AI output revision.

*Verification checklist for Tom:*

☐ Verify whether EcoBot informs Tom that he is interacting with an AI system.

☐ Verify whether EcoBot affects human autonomy by interfering with the Tom's decision-making process in an undesirable way.

☐ Verify that Tom has the power to decide when and how to use EcoBot in any particular situation, including Tom's ability to either decide not to use EcoBot or to override its decision.

The 7-step learning journey, exemplified by the fictional case of Tom and the AI assistant EcoBot, effectively illustrates the practical application of the **TeenTrust-AI** educational framework. Each step is systematically aligned with a core principle of **TeenTrust-AI**, ensuring that educational experiences are ethically guided, developmentally appropriate, and pedagogically robust.

## 5. Conclusion and Future Works

Vulnerable groups such as teenagers and children are consistently more and more exposed to AI and that trend will certainly continue. To guard them for different pitfalls related to usage of AI-powered systems, we need to educate them to assess whether those systems are trustworthy, and therefore, suitable for usage, or not.

This paper presents an educational framework, **TeenTrust-AI**, that uses seven key ALTAI principles to evaluate trustworthiness. While the primary stakeholders of ALTAI principles are AI designers/developers, procurement/legal/compliance officers or specialists, and managers, on a simple use case of an EcoBot, a conversational AI assistant-expert on the topic of climate change, we demonstrate how ALTAI principles can be adapted to a teenage level and successfully applied for trustworthiness verification. The presented approach is topic-agnostic (hence, applicable beyond climate change) and the presented verification checklists can be transferable also to other types of AI systems (other than chatbots).

Future work will evaluate the framework across additional reference topics and reconfigure **TeenTrust-AI** for other types of AI-powered systems (e.g., recommender systems). We will also broaden the target population to include high-school and college students, as well as older adults with

low AI literacy. In the longer term, we plan to conduct longitudinal studies of sustained knowledge, attitudes, and behavior; release open educational resources (OER)[5] and anonymized, ethics-reviewed datasets of annotated student–AI interactions to support replication and establish partnerships with schools and public agencies to align **TeenTrust-AI** with institutional governance and iteratively refine the framework through real-world deployments.

## Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

## References


[1] Y. Hashem, S. Esnaashari, K. Onslow, S. Chakraborty, J. Francis, A. Poletaev, J. Bright, Understanding the impacts of generative ai use on children, Hashem, Y., Esnaashari, S., Onslow, K., Chakraborty, S., Poletaev, A., Francis, J., Bright, J.(2025). Understanding the Impacts of Generative AI Use on Children: WP1 Surveys. The Alan Turing Institute (2025).

[2] D. Kaur, S. Uslu, K. J. Rittichier, A. Durresi, Trustworthy artificial intelligence: A review, ACM Comput. Surv. 55 (2022). URL: https://doi.org/10.1145/3491209. doi:10.1145/3491209.

[3] C. Stix, Artificial intelligence by any other name: a brief history of the conceptualization of "trustworthy artificial intelligence", Discov. Artif. Intell. 2 (2022). URL: https://doi.org/10.1007/s44163-022-00041-5. doi:10.1007/S44163-022-00041-5.

[4] J. Bryson, No one should trust artificial intelligence, Science & Technology: Innovation, Governance, Technology 11 (2018) 14.

[5] X. Shen, Z. Chen, M. Backes, Y. Zhang, In chatgpt we trust? measuring and characterizing the reliability of chatgpt, arXiv preprint arXiv:2304.08979 (2023).

[6] L. Zhong, Z. Wang, Can chatgpt replace stackoverflow? a study on robustness and reliability of large language model code generation, 2024. URL: https://arxiv.org/abs/2308.10335. arXiv:2308.10335.

[7] M. Jang, T. Lukasiewicz, Consistency analysis of chatgpt, 2023. URL: https://doi.org/10.18653/v1/2023.emnlp-main.991. doi:10.18653/V1/2023.EMNLP-MAIN.991.

[8] D. Johnson, et al., Assessing the accuracy and reliability of ai-generated medical responses: an evaluation of the chat-gpt model, Research square (2023).

[9] S. Mitrovic, M. Mazzola, R. Larcher, J. Guzzi, Assessing the trustworthiness of large language models on domain-specific questions, in: M. F. Santos, J. Machado, P. Novais, P. Cortez, P. M. Moreira (Eds.), Progress in Artificial Intelligence - 23rd EPIA Conference on Artificial Intelligence, EPIA 2024, Viana do Castelo, Portugal, September 3-6, 2024, Proceedings, Part III, volume 14969 of *Lecture Notes in Computer Science*, Springer, 2024, pp. 305–317. URL: https://doi.org/10.1007/978-3-031-73503-5_25. doi:10.1007/978-3-031-73503-5_25.

[10] F. Flammini, C. Alcaraz, E. Bellini, S. Marrone, J. Lopez, A. Bondavalli, Towards trustworthy autonomous systems: Taxonomies and future perspectives, IEEE Transactions on Emerging Topics in Computing 12 (2024) 601–614. doi:10.1109/TETC.2022.3227113.

[11] G. Balduzzi, T. Balduzzi, M. Striani, Design an hybrid educational framework for AI ethics in healthcare: Leveraging LLMS and e-learning platforms to empower medical students (full paper), in: E. Marengo, M. Ponticorvo, M. Striani (Eds.), Proceedings of the 1st International Workshop on Education for Artificial Intelligence co-located with the 23rd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2024), Bozen-Bolzano, Italy, November 27, 2024, volume 3902 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024, pp. 116–129. URL: https://ceur-ws.org/Vol-3902/9_paper.pdf.

[12] M. Tanaka, N. Amraan, T. Hurt, E. Greenwald, A. Krakowski, A. Young, M. Cannady, Integrating ethics into AI learning: A socio-technical approach for youth education, in: J. A. Stone, T. T. Yuen,


[5] https://www.unesco.org/en/legal-affairs/recommendation-open-educational-resources-oer

L. Shoop, S. A. Rebelsky, J. Prather (Eds.), Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2, SIGCSE TS 2025, Pittsburgh, PA, USA, 26 February 2025 - 1 March 2025, ACM, 2025, p. 1748. URL: https://doi.org/10.1145/3641555.3705081. doi:10.1145/3641555.3705081.

[13] G. L. Barber, L. Grehan, J. Dunne, B. Davis, H. Qadeer, Towards an educational framework for integrating AI education into second-level education in ireland: Preliminary insights from a national workshop series on AI ethics and privacy (work in progress): Preliminary insights from a national workshop series on AI ethics and privacy (work in progress), in: K. Nolan, R. Monahan, S. Marrone, H. Aldewereld (Eds.), Proceedings of the 2024 Conference on Human Centred Artificial Intelligence - Education and Practice, HCAIep 2024, Naples, Italy, December 2-3, 2024, ACM, 2024, pp. 29–33. URL: https://doi.org/10.1145/3701268.3701275. doi:10.1145/3701268.3701275.

[14] M. Arai, K. Tejima, Y. Yamada, T. Miura, K. Yamashita, C. Kado, R. Shimizu, M. Tatsumi, N. Yanai, G. Hanaoka, REN-A.I.: A video game for AI security education leveraging episodic memory, IEEE Access 12 (2024) 47359–47372. URL: https://doi.org/10.1109/ACCESS.2024.3377699. doi:10.1109/ACCESS.2024.3377699.

[15] N. Bautista, J. Femiani, D. Inclezan, Understanding k-12 teachers' needs for ai education: a survey-based study, in: Proceedings of the Thirty-Ninth AAAI Conference on Artificial Intelligence and Thirty-Seventh Conference on Innovative Applications of Artificial Intelligence and Fifteenth Symposium on Educational Advances in Artificial Intelligence, AAAI'25/IAAI'25/EAAI'25, AAAI Press, 2025. URL: https://doi.org/10.1609/aaai.v39i28.35183. doi:10.1609/aaai.v39i28.35183.

[16] N. Pope, H. Vartiainen, J. Kahila, J. Laru, M. Tedre, A no-code AI education tool for learning AI in K-12 by making machine learning-driven apps, in: Z. Altinay, M. Chang, R. Kuo, A. Tlili (Eds.), IEEE International Conference on Advanced Learning Technologies, ICALT 2024, Nicosia, North Cyprus, July 1-4, 2024, IEEE, 2024, pp. 105–109. URL: https://doi.org/10.1109/ICALT61570.2024.00037. doi:10.1109/ICALT61570.2024.00037.

[17] C. Zhou, D. Kuang, J. Liu, H. Yang, Z. Zhang, A. Mackworth, D. Poole, Aispace2: an interactive visualization tool for learning and teaching artificial intelligence, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, 2020, pp. 13436–13443.

[18] M. Roopaei, R. Roopaei, Gamifying AI education for young minds: The transai adventure in learning, in: IEEE World AI IoT Congress, AIIoT 2024, Seattle, WA, USA, May 29-31, 2024, IEEE, 2024, pp. 129–134. URL: https://doi.org/10.1109/AIIoT61789.2024.10578991. doi:10.1109/AIIOT61789.2024.10578991.

[19] Y. Wang, S. Guo, L. Ling, C. W. Tan, Nemobot: Crafting strategic gaming llm agents for k-12 ai education, in: Proceedings of the Eleventh ACM Conference on Learning@ Scale, 2024, pp. 393–397.

[20] M. Casella, C. Nobile, R. Esposito, M. Luongo, A. Manfredini, N. Milano, O. Gigliotta, D. Marocco, M. Ponticorvo, Not only chatgpt: Educational approaches for embodied artificial intelligence (full paper), in: E. Marengo, M. Ponticorvo, M. Striani (Eds.), Proceedings of the 1st International Workshop on Education for Artificial Intelligence co-located with the 23rd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2024), Bozen-Bolzano, Italy, November 27, 2024, volume 3902 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024, pp. 54–64. URL: https://ceur-ws.org/Vol-3902/11_paper.pdf.

[21] A. Fedele, C. Punzi, S. Tramacere, The ALTAI checklist as a tool to assess ethical and legal implications for a trustworthy AI development in education, Comput. Law Secur. Rev. 53 (2024) 105986. URL: https://doi.org/10.1016/j.clsr.2024.105986. doi:10.1016/J.CLSR.2024.105986.

[22] E. Anagnostopoulou, N. A. Grammatikos, D. Apostolou, G. Mentzas, Trustworthy AI in education: A roadmap for ethical and effective implementation, in: Proceedings of the 13th Hellenic Conference on Artificial Intelligence, SETN 2024, Piraeus, Greece, September 11-13, 2024, ACM, 2024, pp. 25:1–25:7. URL: https://doi.org/10.1145/3688671.3688781. doi:10.1145/3688671.3688781.