20 THE BIASES OF ARTIFICIAL INTELLIGENCE

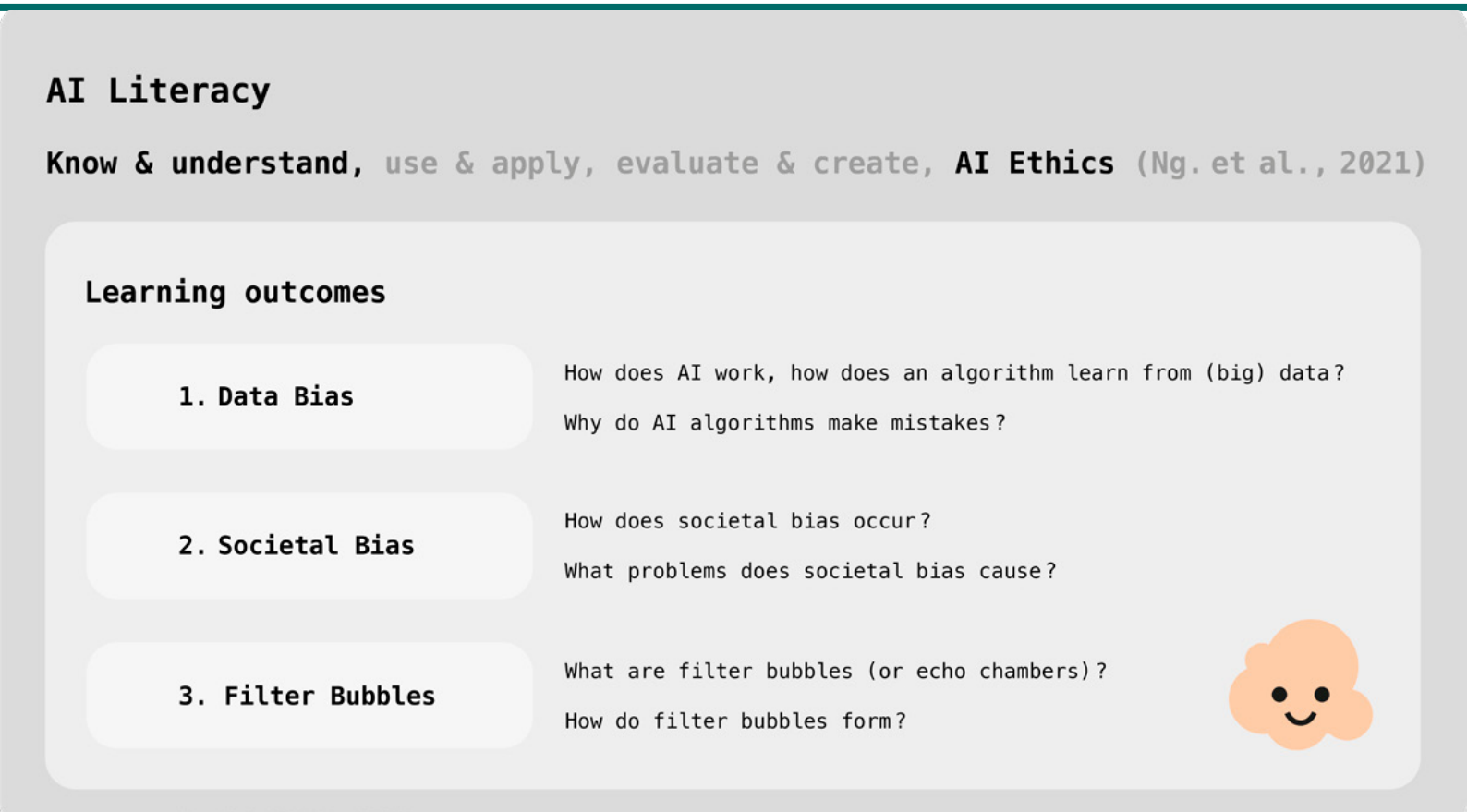
Biljana Petreska von Ritter-Zahony
HEP-VD, University of Teacher Education Vaud, Lausanne, Switzerland
biljana.petreska-von-ritter-zahony@hepl.ch

KEYWORDS
Serious game, artificial intelligence, AI, bias, big data, filter bubbles, echo chambers, high-school, education

ABSTRACT
We have developed a serious game on Artificial Intelligence (AI) for high school students. Due to their recent successes AI-driven processes are pervading a significant number of sectors. However, AI is notoriously difficult to understand because it requires specialised technical and mathematical knowledge.

Our serious game addresses the following learning objectives: How do machines learn? Why are machine learning algorithms biased? How do filter bubbles (or echo chambers) form? The emphasis is on understanding why and how these phenomena occur, not just what they are. We believe that understanding these mechanisms is crucial, as it can empower students to engage in informed debates about artificial intelligence.

Developed in Javascript, the game is easily accessible via a website and can be completed in 20 minutes. It has been tested in a teacher training course and in several high school classes in Switzerland and is highly appreciated (score 5/6, 79 subjects).



CONTEXT

Artificial Intelligence (AI) is a rapidly evolving field of innovation with a deep impact on our society, including education. AI is already challenging assessment practices in education, but it will also redefine the role of teachers and students. We will need to reconsider the competencies that students need to acquire rather than delegate to a machine. There is also a growing need to integrate artificial intelligence into educational curricula (UNICEF, 2021). We have therefore developed a serious game to learn about AI with teachers and students in mind. The game supports AI literacy (Ng. et al., 2021) and encourages the adoption of ethical principles (Adams et al., 2023) by exploring how AI works.

TARGETED ISSUE

The first challenge posed by AI is the exponential nature of its development. This rapid evolution is not reflected in education, where there is a relative lack of knowledge on AI (Luan et al., 2020). Indeed, our data shows that the level of AI literacy among students is generally low.

A second challenge to AI literacy is its technical nature, as the mathematical details that underpin it are difficult to understand. In this work, a lot of effort and thought has gone into making the basic principles and features of AI accessible to everyone, making them as simple and comprehensible as possible. There is a fine line between a simple explanation and a false one, and we have been careful not to cross it.

...the design of a pedagogical sequence that addresses AI at the appropriate cognitive and technical level for high-school students and non-expert adults...

Finally, an important goal was to go beyond a basic understanding of the concepts. For example, the media often write about biased algorithms and report the phenomenon without explaining why it occurs. We believe that in order to understand the bias issue in AI, and why it is impossible to completely overcome it, it is necessary to understand the mechanisms that create bias in the first place.

PROPOSED SOLUTION

We have developed an online game that promotes AI literacy, as students should be conscious users of AI-based systems UNICEF (2021). A direct inspiration for our game was the serious game “The wisdom and/or madness of crowds” by Case (2018), a playful and efficient way to learn about human networks.

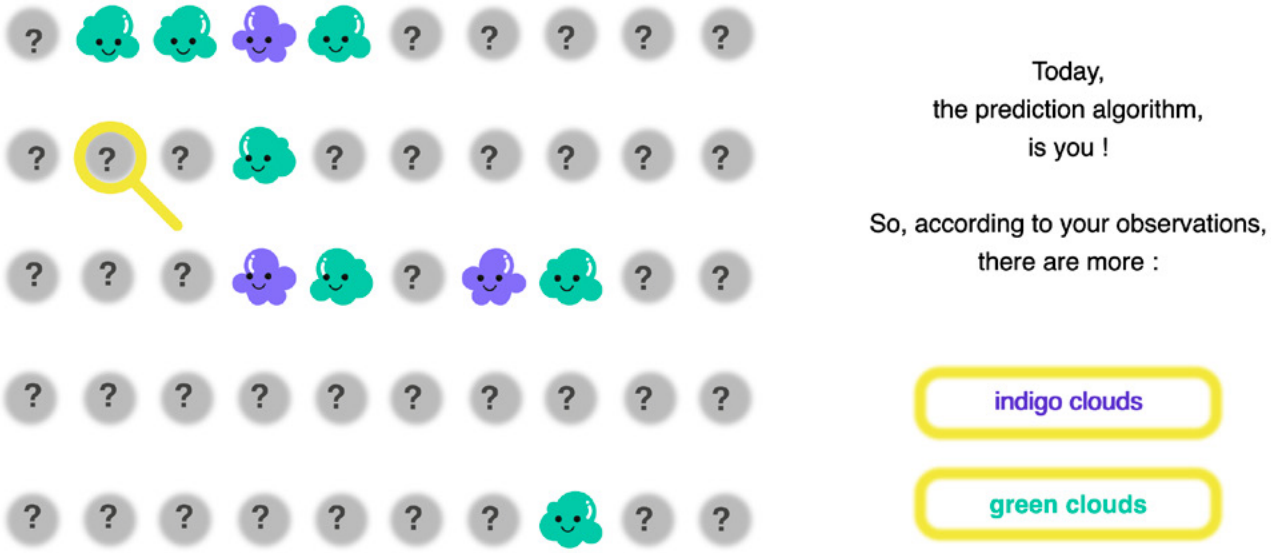
Our first step consisted in identifying the key learning outcomes, which were that students should be able to answer the following questions:

1. Data bias: How does AI work, how does an algorithm learn from (big) data? Why do AI algorithms make mistakes?
2. Societal bias: How does societal bias occur? What problems does societal bias cause?
3. Filter bubbles: What are filter bubbles (or echo chambers)? How do filter bubbles form?

Each learning objective has a dedicated chapter in the game. A final chapter concludes the game and allows players to partially self-assess their understanding of the first and most technical chapter. All chapters can be accessed independently of each other, allowing the teacher to choose the duration and specific content. The game takes approximately 20 minutes to complete.

RELEVANT INNOVATION

Our main contribution is the design of a pedagogical sequence that addresses AI at the appropriate cognitive and technical level for high-school students and non-expert adults (their teachers). We have optimised the learning situations and their explanatory power by manipulating the game interactivity. As the scenario unfolds, the game gives the illusion that the player is in control, while taking the player through a series of fixed stages that progressively construct an understanding of the addressed concepts.



Another contribution is the custom graphics. Special attention was paid to the visual appearance of the characters in the game. The little clouds create a unity in the gameplay, by first representing data and then people. This visual metaphor allows us to address sensitive gender, racial or social issues. The delivered messages are simple, but effective.

We have developed the game in Javascript so that it can be easily accessed online, in a computer science class or at home. Students can progress at their own pace. To accommodate the different speeds of different students we provide many resources that allow them to continue the learning process.

PROJECT OUTCOMES & RESULTS

We first tested the game with ten beta-testers, a mix of AI experts, computer science teachers, social scientists and non-experts. We then provided the game, along with a pre- and post-questionnaire, to several computer science teachers and their students.

The results of the pre-questionnaire show that only a fifth of the 36 computer science teachers were unfamiliar with how artificial intelligence works, and a third had never heard of AI bias. Most had heard of big data, but only a half had heard of filter bubbles. Among students, a fifth of the 145 students said they had some idea of how AI works, and less than a fifth had heard of the term bias in relation to AI. Around a quarter of the students had heard of the term 'big data', but 21 students could describe big data, and only 11 students could give a fairly accurate definition of the term 'filter bubble'.

The post-questionnaire shows that three-quarters of the 56 students, who completed it, felt that they had learned 'a lot' or 'enough' from playing the game, with a fifth still unfamiliar with how AI works. The average score of the game for 79 players is 5 out of 6, with teachers rating the game higher (5.4/6) than their students (4.9/6). One reason for this lower score is that some students felt that they were too guided and wanted more choices in the answers that were offered. Many asked for the game to be more interactive.

CONCLUSION

This serious game tries to make the mechanisms of artificial intelligence and its problems more tangible for high school students. Looking at some of the students' responses to the post-questionnaire, we are confident that the game is at the appropriate level. Most of the students appreciated the game and said that it helped them learn about AI.

PERSPECTIVES & NEEDS

An open question is whether the game is also accessible to 15 year olds or younger. We need to address some important criticisms in the future, such as better interactivity. We hope that this game will contribute to raising awareness about AI through education.

ACKNOWLEDGEMENTS

MODULO: Projet d'éducation numérique du canton de Vaud.

https://modulo-info.ch/

REFERENCES

- Adams, C., Pente, P., Lemermeyer, G., Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. Computers and Education: Artificial Intelligence, Volume 4.
- Case, N. (2018). The Wisdom and/or Madness of Crowds.
- https://ncase.me/crowds/
- Luan, H., Geczy, P., Lai, H., Gobert, J., Yang. S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., Tsai C.-C. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. Frontiers in Psychology, Volume 11. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.580820/full
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review, Computers and Education: Artificial Intelligence, Volume 2.
- UNICEF (2021). Policy guidance on AI for children 2.0. United Nations Children's Fund. https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-po licy-guidance-AI-children-2.0-2021.pdf.pdf.
- * https://mi.hepl.ch/projects/ia/ia.html [original game in French] https://mi.hepl.ch/projects/ia/bias.html [in English]