Democratizing AI through Human-Centered AI Literacy Practices

CHINASA T. OKOLO, The Brookings Institution, United States

The increasing usage of AI in high-stakes decision-making underscores a pressing need for various stakeholders to understand AI, learn how to identify AI-generated content, and become aware of its societal risks. This project aims to answer the research question, "How can we leverage human-centered approaches to build AI literacy in 9-12 students?" by detailing outcomes from engaging underrepresented secondary school students in a 5-day workshop series consisting of brief lectures, hands-on activities, and short research assignments. We find that the workshop improved students' knowledge about AI and the ethical implications of using these technologies. Our work proposes an actionable framework to advance AI literacy upskilling efforts and has developed the workshop content into an open-source AI literacy curriculum.

CCS Concepts: • **Applied computing** → **Education**; • **Computing methodologies** → **Artificial intelligence**; *Machine learning*; • **Human-centered computing**;

Additional Key Words and Phrases: AI literacy, computing education, equity, human-centered AI, generative AI, technology ethics, AI pedagogy

1 INTRODUCTION

For the first time in recent history, artificial intelligence (AI) has become a public conversation. From ChatGPT, which sparked the "AI revolution," to image generators like Stable Diffusion and DALL-E, people worldwide have begun actively engaging with AI. Despite this, many people are still passive consumers and sometimes unwilling targets of these systems. Additionally, despite researchers highlighting issues of bias in AI, many of its risks are still not understood [9]. This lack of understanding could exacerbate discrimination and misinformation from AI systems, harming users. As AI development is primarily centralized in the hands of powerful tech companies in the West [19], it will also be essential to understand how to engage marginalized populations to become equitable stakeholders in AI. Such efforts could empower these users to understand the impacts of AI, regain agency from automated decision-making, and potentially seek recourse from harms committed by these technologies.

Given the ability of AI to make essential decisions in fields such as healthcare, agriculture, and education, there is a pressing need for various stakeholders to understand AI, learn how to identify AI-generated content, and become aware of its risks to society. Improving AI literacy is just one way to do this. Researchers have defined AI literacy as *"a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace"* [14]. While there exist no prominent AI literacy frameworks to aid in facilitating pedagogy, there is an emerging area of work focused on developing AI literacy in primary and secondary school educational contexts [13, 15, 17, 26, 27].

To ensure that various stakeholders understand the benefits, limitations, and ethical implications of AI, much investment is needed to support AI literacy efforts. Creating the social and technical infrastructure to support AI literacy will serve public interests and shift the democratization of AI to include novice technology users. To progress toward this goal, the research team conducted a 5-day workshop series with underrepresented secondary school students in a rural county in New York State to engage them in various aspects of AI, ranging from data collection to algorithmic development to the ethical considerations of generative AI. We found that incorporating critical reflection assignments

Author's address: Chinasa T. Okolo, Center for Technology Innovation, The Brookings Institution, 1775 Massachusetts Ave, Washington, D.C., United States, cokolo@brookings.edu.

along with short lectures, videos, and experimentation with AI tools allowed students to understand their role as consumers of AI and how they can thoughtfully engage with these systems.

2 DESCRIPTION OF PRACTICE

To answer the research question that guided the goals of this workshop series, I conducted five (5) workshop sessions across two weeks in July 2023. This section details the participants, workshop curricula, and general findings from engaging students.

2.1 Participants

To pilot this work, I recruited a group of local high school students through STEP-UP, a summer internship program directed by graduate students at Cornell University. The program aims to positively impact the high school student population in the university area through outcome-based programming, scientific exposure, diversity initiatives, and interactions with college and graduate students. Each of the four workshop participants identified as female. One of the participants identified as Black, one as mixed-race, and two as Asian. One of the students was entering the final year of middle school (8th grade), one was entering high school (9th grade), and the other two had completed one year of high school.

2.2 Workshop Goals

To help guide the implementation of the workshop series, I defined several goals:

- Students participating in this 5-day workshop series will learn about various aspects of AI, including the socio-ethical implications of emerging AI methods such as generative AI.
- Students will critically engage with these methods and better understand how to use AI responsibly.
- Students will also improve their AI literacy by learning to recognize AI-generated content and understanding the shortcomings of AI tools like large language models and text-to-image models.

2.3 Workshops

Throughout five 3-hour sessions, workshop participants engaged in hands-on activities, brief lectures, and short research assignments. Below, I detail the aims of the workshops and the content provided in each session. The primary author has developed formal lesson plans and lecture slides from the workshops to be shared open-source.

2.3.1 Session I: General Introduction to AI. The first session of this workshop series provided students with a general introduction to AI. The session began by engaging students in a group discussion, where they shared their knowledge of AI and aspects they were interested in learning more about. To guide this discussion, students were asked three questions:

- What do you know about AI?
- What do you want to know about AI?
- What do you think is important for students your age to know about AI?

Then, students were introduced to the AI concepts through a short lecture. First, the lecture provided definitions of artificial intelligence and machine learning, illustrating the relationship between these two fields. It also provided examples of other subfields of AI, like deep learning, computer vision, and natural language processing. Next, the lecture detailed the history of AI, starting from work done in the 1950s by Alan Turing and leading to present-day advancements

in generative models and smart assistants. It also provided example use cases of AI, so students could understand how these methods are applied in real life. Finally, the lecture also briefly introduced students to the implications of AI, which was then followed up more thoroughly in Session III.

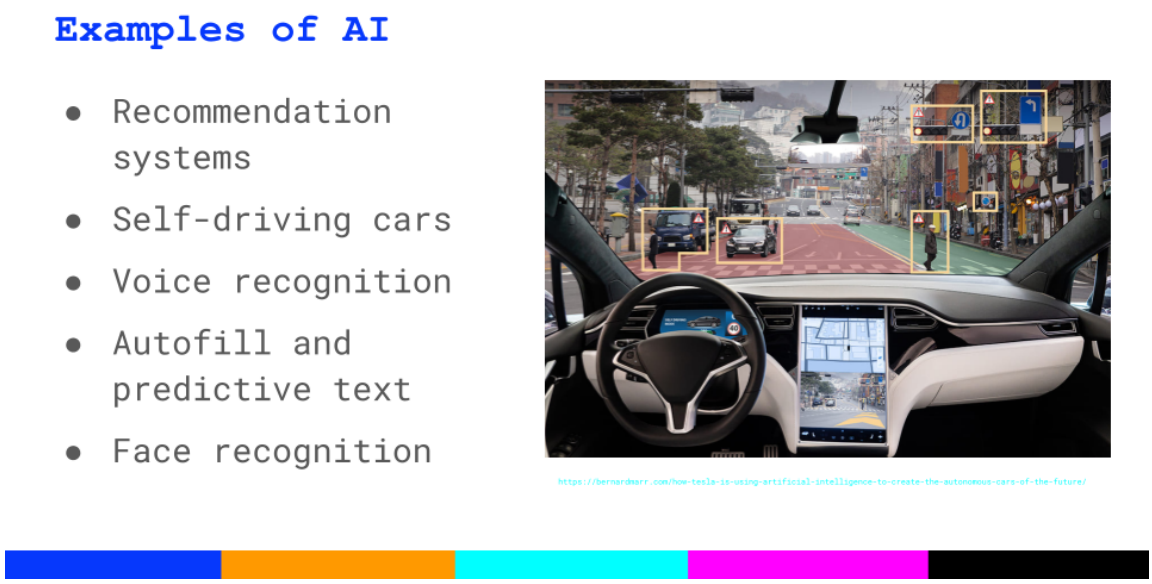

Fig. 1. A lecture slide from Workshop Session I detailing example applications of AI in real-world contexts.

AI Activity To give students first-hand experience with AI, the next section of the first workshop incorporated the "Discover AI in Daily Life" lesson from Google [24]. This activity involved (1) students choosing three words and translating them using Google Translate into a language of their choosing, (2) using Google AutoDraw to create an AI-generated image of each word, and (3) creating a Google Slides presentation with these words and AI-generated images. After presenting their slides individually, we debriefed on the activity and transitioned into a brainstorming session where students discussed the following questions:

- How do you use AI in your daily life?
- What are some interesting things you could do with AI?
- What are some potential use cases of AI?

2.3.2 Session II: Data and Algorithms. In the second workshop session, students learned about the importance of data and algorithms in machine learning (ML). The session began with a group discussion on their knowledge of these concepts. To guide the discussion, students were asked the following questions:

- What do you know about datasets?
- What do you know about algorithms?
- Do you know how they impact ML models?

Then, students were introduced to the concepts of data, algorithms, and model training through a brief lecture. First, the lecture defined data and datasets, providing examples of popular datasets used in machine learning. It then defined algorithms and provided examples of what algorithms can be used for. To lead up to the full lecture on algorithmic bias, which would take place in Session III, students were introduced to the topics of dataset and algorithmic bias and watched a short video explaining how bias impacts ML models [7]. It then explained how data is used in machine learning, reviewing the collection, cleaning, and labeling processes. The lecture briefly ended by explaining how models are trained and evaluated.

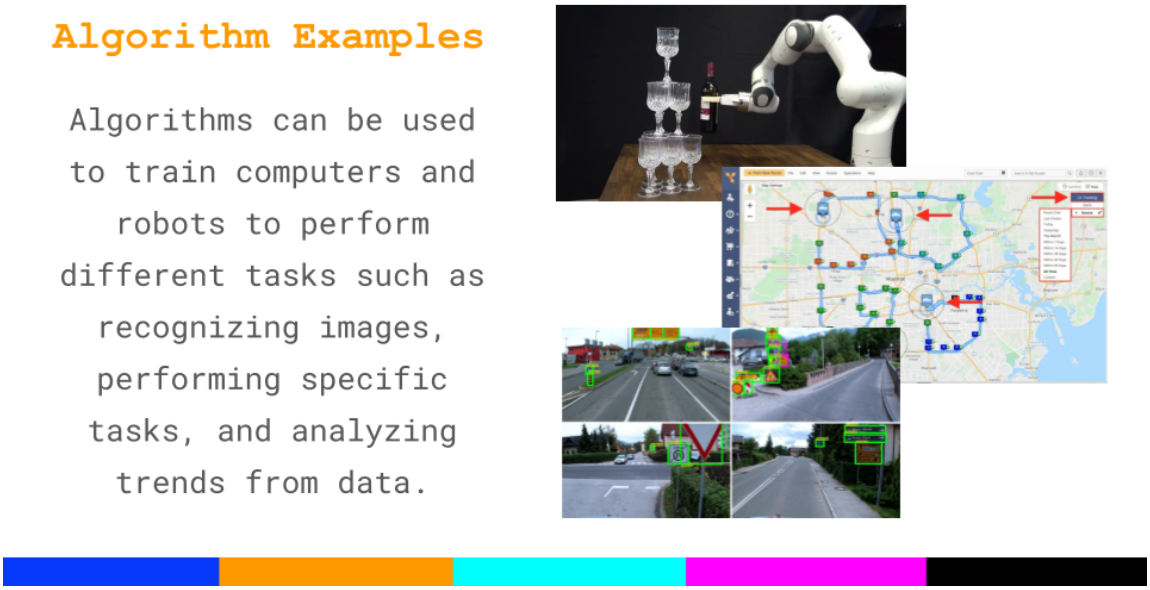

Fig. 2. A lecture slide from Workshop Session II detailing examples and functions of algorithms.

ML Activity. Next, students participated in a hands-on activity to create a simple ML model using the "Teachable Machines" platform developed by Google [8]. In this activity, students gathered data of their choosing (images, sounds, poses, etc.), trained a machine learning model, exported it, and wrote a short response to their experimentation. Students were directed to test what data their models work on, test what data their models don't work well on, strategies they used to "trick" their trained models, and how tweaking the epochs, batch size, and learning rate improves model performance. After students presented their reports, we ended the session by debriefing on the activity.

2.3.3 Session III: Algorithmic Bias. In the third workshop session, students learned about the ethical implications of AI. Similar to the previous sessions, I began the session with a group discussion on their existing knowledge of the session topic, in this case, covering bias and its impact on ML models. To guide the discussion, students were asked the following questions:

- What do you know about bias?
- What are some examples of bias?

- Do you know how bias impacts ML models?

Next, a brief lecture detailed information about bias and case studies regarding the ethical implications of AI. First, I defined bias and presented various types of bias, including confirmation, stereotyping, availability, implicit, anchoring, and in-group bias. Next, I specifically presented topics relating to bias in machine learning, covering dataset bias and algorithmic bias. The lecture also incorporated two separate videos to help students further visualize these concepts [6, 7].

Fig. 3. A lecture slide from Workshop Session III detailing algorithmic bias.

Bias Case Study. To lead into the activity for this workshop session, students watched a video on algorithmic bias that detailed how bias shows up in search results [1]. In this activity, students were instructed to find examples of AI bias incidents covered in the media or in research articles and create a short write-up on one example. To guide their writing process, I asked students to detail *what* their bias case study is about and what field it impacts (medicine, finance, hiring, etc.), *who* (company, researchers, etc.) developed the biased tool or algorithm, *when* the bias incident occurred, *how* it affected users, and *why* they think it is an important issue. After students presented their case studies, we moved on to a short lecture about the ethical issues of data annotation work, presenting a video on data labeling [23] and asking students to read a recent article on workers involved in labeling data for ChatGPT [20]. The session ended with a debriefing on the activity.

2.3.4 Session IV: Introduction to Generative AI. In the fourth workshop session, students learned about the basics of Generative AI. To begin this session, I initiated a group discussion on the students' knowledge of generative AI. To guide the discussion, students were asked the following questions:

- What do you know about generative AI?

- What are some examples of generative AI?

Then, students were introduced to the concepts of generative AI in a lecture. First, I defined generative AI and shared a timeline of developments in generative models from the 1950s to the present. Next, I shared examples of how generative AI is being used (image generation, question answering, music composition, etc.) and transitioned to the next topic of large language models (LLMs) and text-to-image models. After defining LLMs, the instructor detailed how they function and shared examples of popular models like ChatGPT, Bard, Midjourney, and Dall-E. To codify these concepts, students watched a short video from The Economist [5]. The last section of the lecture discussed the limitations of generative AI and examples of gender, cultural, and regional bias produced by these systems. To transition into the activity for this session, students read an article on gender bias from AI text-to-image generators [18] and watched a video from Bloomberg detailing a recent study on racial and gender bias in generative AI [25].

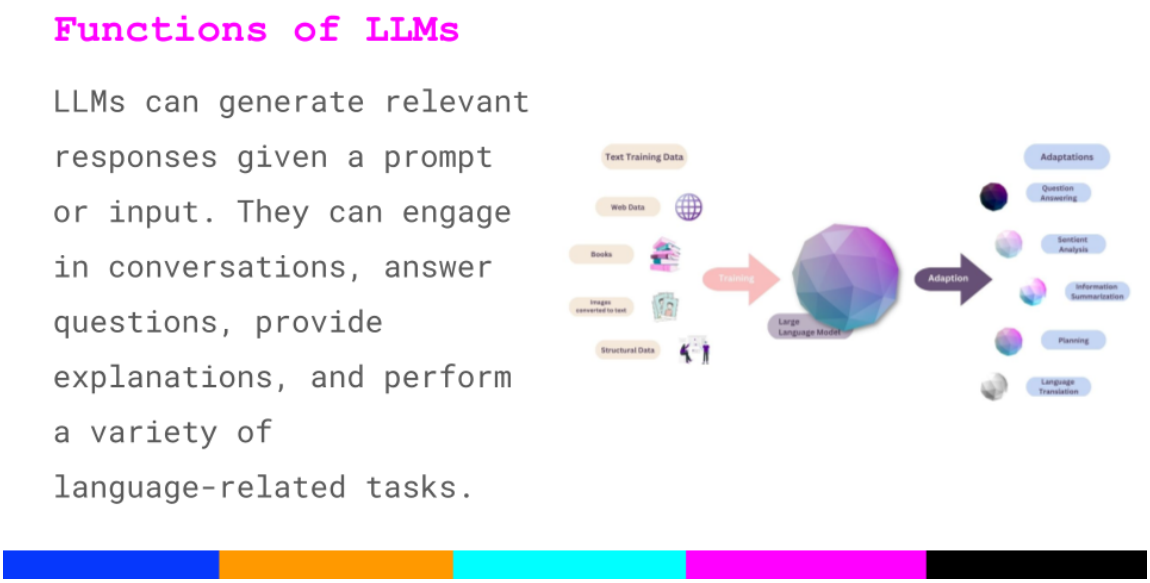

Fig. 4. A lecture slide from Workshop Session IV detailing functions of large language models (LLMs).

Generative AI Activity. To give students more context into how generative AI models work, the first part of the activity for this session involved experimentation with a tool developed by the New York Times to detail the process of training an LLM "from scratch" [3]. Students could select literature from Jane Austen, Shakespeare, Moby-Dick, Star Trek, Harry Potter, and the Federalist Papers and go through an interactive demonstration to learn how LLMs "learn" language from 250 to 30,000 rounds of training on these texts. Next, to help students develop strategies to discern between original and AI-generated content, students engaged with an interactive quiz developed by the New York Times to test if they could differentiate if specific text samples were written by a fourth-grade student or AI-generated. For the last part of this activity, students played an interactive game to understand how LLMs can be manipulated to reveal information through a technique known as "prompt injection", where users manipulate LLMs to perform

unintended actions involuntarily [22]. For each section of this activity, students were asked to write up their experiences interacting with these tools and gave short presentations. The session ended with a debriefing on the activity.

2.3.5 Session V: Implications of Generative AI. In the fifth and final workshop session, students reviewed concepts about Generative AI and experimented with popular AI tools to generate images and text. To begin this session, I initiated a group discussion on the students' experiences using generative AI tools. To guide the discussion, students were asked the following questions:

- What do you know about ChatGPT or other generative AI tools?
- Have you experimented with generative AI tools?
- How was your experience with these tools?

Next, I reviewed the concept of generative AI by watching the first ten minutes of a video developed by Google [4] and transitioned into the planned activity.

Ethical Implications of Generative AI Activity. To give students an opportunity to increase their direct engagement with generative AI tools, students experimented with ChatGPT and Microsoft Image Generator. In this experimentation, students generated text and image content and wrote up a short reflection analyzing what these tools do well and the kinds of tasks they appear to perform poorly at. Students then presented their findings. Next, the students watched a video on the ethical implications of ChatGPT [16] and read an article from the New York Times on university responses to AI chatbots [11]. After engaging with this content, students were asked to create a short write-up guided by the following questions:

- Why do you think many educators are worried about ChatGPT and similar technologies?
- Who is responsible for ensuring these tools are used responsibly?
- When should Generative AI not be used?
- How do you think students can misuse ChatGPT and other technologies?
- Why is it important to use generative AI responsibly?

To gain further perspective on the ethical concerns of using generative AI tools in the classroom, students watched a video detailing how a school in Hong Kong embraced this technology [21]. Finally, to incorporate their learnings from the workshop sessions, I engaged students in a discussion about the implications of using generative AI and how these methods could be ethically used to complement their learning. The last session ended by asking students to fill out a post-workshop survey, debriefing on the workshop, and providing opportunities for the students to ask any lingering questions.

2.4 Analysis

Pre-Workshop Survey. To guide the development of the workshop materials, students were asked to fill out a pre-workshop survey before the start of the first session. This survey asked the students what they knew about AI, if they had used any AI tools, their understanding of AI (using a Likert scale rating from 1-5, with 1 indicating low AI understanding and 5 indicating high AI understanding), their familiarity with the ethical implications of AI, and what they were specifically interested in learning about AI.

When asked about their AI knowledge, all of the students generally knew that "AI" stood for "artificial intelligence," with one student in particular mentioning that it aims to "imitate human knowledge and can generate an answer to almost anything you ask it." Students often mentioned their knowledge of AI in relation to text-generation applications

like ChatGPT. When rating their understanding of how AI works, the students had an average rating of 3 with a median rating of 3.5. Exactly half of the students responded that they were familiar with the ethical implications of AI. When asked what they were specifically interested in learning about AI, students mentioned wanting to understand how AI functions, examples of real-world applications of AI, the origin of AI, and the potential implications of this technology.

Post-Workshop Survey. To understand the impact of the workshop on their perceptions of AI, students also filled out a post-workshop survey. This survey asked them what they learned about AI, their interest in using AI tools in the future based on their workshop experiences, how the workshop impacted their understanding of AI and its associated ethical implications, and any remaining questions they had about AI.

When asked about their takeaways from the workshop, students mentioned that they learned how AI is developed, what it can be used for, and its ethical implications. Many of the students emphasized their improved perceptions of AI, especially as it pertains to the limited ability of AI to become sentient or "take over the world." Students also recognized that AI is increasingly being used in various aspects of their lives and that there are many implications to using AI. All of the students responded positively to using AI tools in the future. When rating if their understanding of how AI works improved (also scored using a Likert scale rating from 1-5, with one indicating "strongly disagree" and five indicating "strongly agree"), students provided an average rating of 4.75 with a median rating of 4.5.

2.5 Outcomes

We found unique insights from engaging the students in this 5-day workshop series. First, while all of the students currently engaged with AI tools such as ChatGPT, all of them had a limited understanding of AI concepts (i.e., recommendation systems, data labeling, large language models), which was not unexpected given their respective educational levels. However, some of the students were somewhat familiar with the implications of AI, from exposure to personal readings or popular news topics. The workshop participants had various interests in understanding the inner workings of AI. For example, one student mentioned wanting to know what *"the code [for AI models] actually looks like"* and another wanted to understand how to actually build AI models, aside from what they had experimented with through Teachable Machine. This workshop also provided an opportunity to expose students to the harms of AI, especially as it pertains to the labor involved in the production and annotation of data, an essential aspect of training AI systems. From this exposure, critically engaging students while allowing them to reflect on their personal usage of AI tools improved their understanding of AI and their ability to recognize potential harms.

3 POSITIONALITY

The workshop team is based in the United States and has developed curricula for and conducted fieldwork with underserved communities in low-resource regions within the United States and the Global South. The primary author identifies as female and is from an underrepresented racial background (in the context of the United States). The author also has 8+ years of experience teaching computing and AI topics to students ranging from kindergarten to undergraduate levels. As an AI researcher whose work centers on marginalized communities, I believe in elevating the voices of local populations and actively including them in AI development. I approach my research through an equity-driven [10] and emancipatory action [2, 12] mindset, where I aim to identify the opportunities and challenges of improving AI literacy in marginalized populations while underscoring the needs of those who will interact with or be subjects of AI-enabled technologies.

4 LIMITATIONS AND ASSUMPTIONS

This project is subject to several potential limitations. The primary constraint of this project lies in our small participant sample size, as the program the workshop series was a part of intentionally targeted a small number of students for each workshop. Additionally, all the students who participated in the workshop were female, so our preliminary findings may not generalize across students of different genders. Given that there are no standardized methods for measuring AI literacy, I also acknowledge that students may have overestimated their AI knowledge in the pre-workshop survey. We also found that some experimentation was needed to ensure I had enough content for the sessions because they were longer than traditional lectures (e.g., 3 hours per workshop vs. 1-hour class periods). The first workshop session was primarily lecture-based and included one activity, which wasn't long enough for 3 hours. In subsequent sessions, I developed slightly longer lectures incorporating videos and then leveraged article readings combined with short group discussions on the assigned articles before starting the primary activity. If educators are interested in adapting these workshops to a larger group of students, it may be necessary to explore different measures to gauge students' initial AI knowledge and track its change over time, which could help them choose what content to include for each workshop session. Additionally, with the short lead time to produce the content for the workshops, the team found it difficult to obtain short video explainers on more complicated topics discussed in the workshop, such as generative AI, that were accessible to the age group of the participants. To get around these constraints, I often found videos geared to more technical audiences and shortened the amount of content I would show students. For example, in the last session, I showed students the first 10 minutes of a video on Generative AI [4] to briefly recap the more in-depth introduction provided to students in the previous workshop. The workshop series described in this report primarily focused on teaching secondary school students, so some of the content provided in the lecture slides may not be accessible to younger students. However, considering the urgent need for AI literacy to reach a more significant number of other key stakeholders, working with teachers and students at primary and tertiary schooling levels is an essential area for future work.

5 PROJECT DELIVERABLES

To understand how to leverage the findings from this work to inform broader efforts on AI literacy, we developed a preliminary framework to characterize the potential outcomes of AI upskilling (outlined in Figure 5). We posit that AI literacy should incorporate versatile learning to ensure learners are exposed to primary, intermediate, and advanced AI concepts. From this exposure, students benefit from enhanced awareness of the limitations, benefits, and harms of AI, thus enabling students to become responsible stewards of AI who use these tools ethically and are comfortable relying on their respective intuition to make high-stakes decisions.

The content from these workshop sessions was adapted into an open-source curriculum, "Beyond AI Hype", available online. The lesson plans are linked below:

- Session I Lesson Plan
- Session II Lesson Plan
- Session III Lesson Plan
- Session IV Lesson Plan
- Session V Lesson Plan

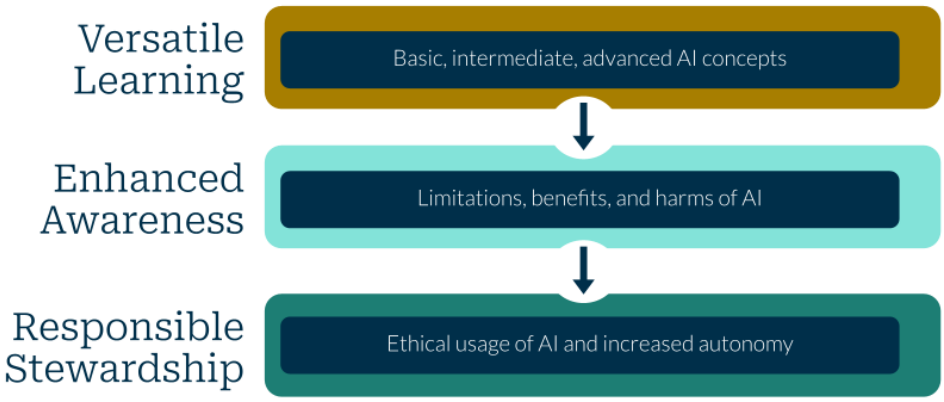

Fig. 5. The proposed AI literacy framework.

6 FUTURE DIRECTIONS

The pilot of this 5-day workshop provides various avenues for future work. Future steps for this research involve scaling up the workshops to reach more students and gain feedback (e.g., through structured surveys) on how AI curriculum could be enhanced for various age groups. While the planning and implementation of this workshop were not incorporated into a formal research study, I also plan to examine how to leverage interviews, focus groups, and co-design workshops to design relevant AI literacy curricula. We also find it necessary to make progress towards the development of standardized methods for measuring AI literacy. Understanding how to implement effective evaluation methodologies for measuring AI literacy will help provide tangible metrics that future studies can incorporate.

We also find it essential to gain perspectives from educators, and future avenues for this work could involve co-designing with K-12 educators to develop AI curricula for students across primary and secondary educational settings. To help this process, immediate future steps for this work will involve examining contemporary measures that quantify AI knowledge and reviewing existing work in HCI and computer science education to inform the development of a human-centered AI framework for building AI literacy. This AI literacy framework would then guide AI curricula co-design, contributing to future efforts to advance AI literacy.

7 CONCLUSION

This experience report details a 5-day workshop series conducted with underrepresented middle and high school students in Upstate New York. We find that integrating critical reflection assignments along with short lectures, videos, and experimentation with AI tools into the workshop sessions provides students an opportunity to understand their role as consumers of AI and ways they can thoughtfully engage with these systems. Our work contributes to existing

literature on AI literacy and computing education while also providing additional resources that educators can leverage to teach AI concepts and engage students in emerging AI topics. The increasing interest in building AI literacy in varying stakeholder populations outside of secondary school settings potentially suggests that our findings could be relevant to a broader set of populations. However, future curricula development and implementation will be required to understand how our work can generalize to other domains.

ACKNOWLEDGMENTS

We are thankful to all of the participants who contributed valuable insights to our study. This work was funded by the ICICLE Educational Fellows program.

REFERENCES

[1] Algorithms and Society. 2020. Algorithmic Bias, Discover How It's Affecting You! Retrieved July 1, 2023 from https://www.youtube.com/watch?v=7SVin0_ePGQ

[2] Jeffrey Bardzell and Shaowen Bardzell. 2015. What is Humanistic HCI? In *Humanistic HCI*. Springer, 13–32.

[3] Aatish Bhatia. 2023. Let Us Show You How GPT Works — Using Jane Austen. *The New York Times* (2023). Retrieved July 1, 2023 from https://www.nytimes.com/interactive/2023/04/26/upshot/gpt-from-scratch.html

[4] Google Cloud. 2023. Introduction to Generative AI. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=G2fqAlgmoPo&t=1s

[5] The Economist. 2023. Generative AI: what is it good for? Retrieved July 1, 2023 from https://www.youtube.com/watch?v=gCDacaohqaA

[6] Institute for Public Policy Research. 2020. A Algorithmic Bias Explained. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=tia5OHE98F4

[7] Google. 2017. Machine Learning and Human Bias. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=59bMh59JQDo

[8] Google. 2017. Teachable Machine. Retrieved July 1, 2023 from https://teachablemachine.withgoogle.com/

[9] Alexa Hagerty and Igor Rubinov. 2019. Global AI ethics: a review of the social impacts and ethical implications of artificial intelligence. *arXiv preprint arXiv:1907.07892* (2019).

[10] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing Community-based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–25.

[11] Kalley Huang. 2023. Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach. *The New York Times* (2023). Retrieved July 1, 2023 from https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html

[12] Os Keyes, Josephine Hoy, and Margaret Drouhard. 2019. Human-computer Insurrection: Notes on an Anarchist HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[13] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM technical symposium on computer science education*. 191–197.

[14] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–16.

[15] Radu Mariescu-Istodor and Ilkka Jormanainen. 2019. Machine learning for high school students. In *Proceedings of the 19th Koli calling international conference on computing education research*. 1–9.

[16] NBC News. 2023. Breaking down ethical questions surrounding new chatbot GPT-4. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=C02-YMqwpS8

[17] Davy Tsz Kit Ng, Jiahong Su, Jac Ka Lok Leung, and Samuel Kai Wah Chu. 2023. Artificial intelligence (AI) literacy education in secondary schools: a review. *Interactive Learning Environments* (2023), 1–21.

[18] Kristina Nikolic and Jelena Jovicic. 2023. Reproducing inequality: How AI image generators show biases against women in STEM. Retrieved July 1, 2023 from https://www.undp.org/serbia/blog/reproducing-inequality-how-ai-image-generators-show-biases-against-women-stem

[19] Chinasa T Okolo. 2023. Addressing global inequity in AI development. In *Handbook of Critical Studies of Artificial Intelligence*. Edward Elgar Publishing, 378–389.

[20] Billy Perrigo. 2023. OpenAI Used Kenyan Workers on Less Than $2 Per Hour. *TIME* (2023). Retrieved July 1, 2023 from https://time.com/6247678/openai-chatgpt-kenya-workers/

[21] South China Morning Post. 2023. How a Hong Kong school embraces ChatGPT in the classroom. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=z3yYmuH0Ymw

[22] Deval Shah. 2023. The ELI5 Guide to Prompt Injection: Techniques, Prevention Methods Tools. *Lakera AI* (2023). Retrieved August 1, 2023 from https://www.lakera.ai/blog/guide-to-prompt-injection

[23] DW Shift. 2023. Doing Grueling Work for an AI: Data Labeling. Retrieved July 1, 2023 from https://www.youtube.com/watch?v=ug_p2wHhla0&t=1s

[24] Google Applied Digital Skills. 2022. Discover AI in Daily Life. Retrieved July 1, 2023 from https://applieddigitalskills.withgoogle.com/c/middle-and-high-school/en/discover-ai-in-daily-life/overview.html

[25] Bloomberg Technology. 2023. Calling Out the Racial and Gender Bias in Generative AI. Retrieved July 1, 2023 from https://www.bloomberg.com/news/videos/2023-06-09/calling-out-the-racial-and-gender-bias-in-generative-ai-video

[26] Randi Williams, Stephen P Kaputsos, and Cynthia Breazeal. 2021. Teacher perspectives on how to train your robot: A middle school AI and ethics curriculum. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15678–15686.

[27] Helen Zhang, Irene Lee, Safinah Ali, Daniella DiPaola, Yihong Cheng, and Cynthia Breazeal. 2023. Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education* 33, 2 (2023), 290–324.