BOARD # 75: One Teacher's Experience Adapting an Innovative, Flexible Computer Vision Curriculum in a Middle School Science Classroom

Dr. Christine Wusylko, University of Florida

Christine a postdoctoral fellow at the University of Florida. She draws on over 10 years of experience teaching science and technology across grade levels K-16, to produce useful and usable knowledge, which is both driven by problems of practice and is theoretically grounded. Her research and development program is centered on helping young people develop AI and STEM literacy in authentic learning environments.

Rachel Still, University of Florida
Dr. Pavlo Antonenko

Pavlo "Pasha" Antonenko is an Associate Professor of Educational Technology at the University of Florida. His interests focus on the design of technology-enhanced learning environments and rigorous mixed-method research on the effective conditions for tec

Brian Abramowitz, University of Florida

Brian Abramowitz is a coordinator for the Scientist in Every Florida School program at the University of Florida. His research interests center around the implementation of AI in K-12 classrooms.

Dr. Jeremy A. Magruder Waisome, University of Florida

Dr. Jeremy A. Magruder Waisome is the Thomas O. Hunter Rising Star Assistant Professor in the Engineering Education Department at the University of Florida (UF). Her research focuses on self-efficacy and critical mentorship in engineering and computing. She is passionate about broadening participation and leveraging evidence-based approaches to improve the engineering education environment for minoritized individuals.

Victor Perez
STEPHANIE KILLINGSWORTH, University of Florida

WIP: One Teacher's Experience Adapting an Innovative, Flexible Computer Vision Curriculum in a Middle School Science Classroom

Introduction

Artificial intelligence (AI) is predicted to be one of the most disruptive technologies in the 21st century [1], and to prepare all young people to live and work in an AI infused world, many are calling on teaching Computer Science (CS) and AI across grade bands [2]. This is important even if children will not have a disciplinary preference for AI, as being introduced to AI and developing positive attitudes towards CS are necessary steps to developing educated citizens who will make responsible and knowledgeable decisions about science related issues [3], [4], [5].

Integrating AI into the K-12 curriculum, however, is a challenging task. While there are guidelines and emerging frameworks for integrating AI (e.g., [6]), there are currently no national standards for teaching AI in K-12. Further, teachers may not see clear connections between AI and K-12 core standards as GenAI is a rapidly emerging technology and inservice and preservice teachers alike have had little training in this topic [7].

To help teachers integrate AI into their curriculum, developers and educational researchers are creating AI educational innovations for young people, many of which introduce AI using large language models (LLMs) and chatbots (e.g. [8], [9]) in afterschool settings (e.g. [10], [11]). While these endeavors have been successfully implemented, there remains a gap for introducing AI technologies beyond LLMs and chatbots in the formal K-12 setting. Specifically, computer vision is an underutilized and accessible way to introduce young people to CS and AI, and has potential to be integrated into core middle school science standards.

To address this gap, our interdisciplinary team of paleontologists, natural history museum educators, computer science engineers, and educational technology researchers designed and developed an innovative, flexible computer vision curriculum for middle school science classrooms. The curriculum called Shark AI is funded by the NSF (award # 2147625) and blends CS and paleontology as middle school students build and evaluate their own computer vision model used to classify fossil shark teeth. The curriculum consists of five flexible modules aligned with national and Florida science standards. They also address the five core areas of AI, according to the AI4K12 framework for AI education in K-12 schools. These modules include an introduction to AI, classifying shark teeth and data, training and evaluating machine learning models, identifying bias in datasets, and building a machine learning model.

Shark AI is implemented as a three-year project, and is currently in year three of implementation. To date, we have partnered with 57 teachers and 597 students in 47 schools. This work-in-progress paper is part of an ongoing investigation to explore how teachers adapt this curriculum to their own classroom, and describe successes and lessons learned, which can be applied to others wishing to design and develop similar curricula. In this paper, we focus on the Shark AI implementation of one 7th grade science teacher and describe the implementation choices and changes she made for her own classroom and students' needs, including her modifications of the curriculum, researcher observations, student achievement pre and post tests, analysis of student work, and students' science related attitudes before and after the modules.

Conceptual Framework

The "5 Big Ideas in AI" created by the Artificial Intelligence (AI) for K-12 initiative (AI4K12) guided the development of both the Shark AI curriculum and the pre and post learning assessments for each module [12]. The framework outlines five major concepts around AI and is used nationally for teaching AI in a K-12 setting. These "5 Big Ideas in AI" are perception, representation and reasoning, learning, natural interaction, and societal impact (Fig. 1). The following introduces each big idea and each connected module and learning activity. Appendix A details the learning assessment items for each module and the connection to each big idea.

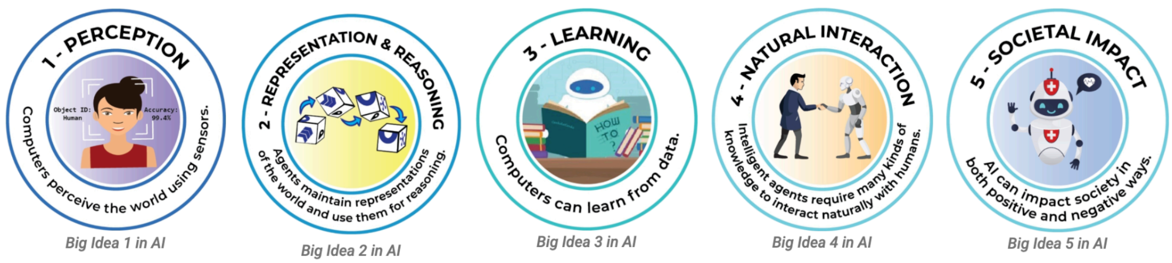

Fig. 1. The "5 Big Ideas in AI" [13]

Perception

Computers use input devices and sensors, such as cameras, to perceive the world. In Module 1 of Shark AI curriculum, students are introduced to the different ways computers can perceive the world. The remaining Modules 2-5 guide participants in creating and evaluating computer vision models created with the tool Google Teachable Machine (GTM). In these learning activities, students explore how computer vision models perceive the world and make predictions based on that data.

Representation & Reasoning

Computers construct representations of the world using data and reason with those representations to create new information. In Modules 2-4 students are tasked with creating datasets of shark teeth with enough quality images to create a functioning model. Through guided activities, students test the accuracy of their model and critically evaluate the datasets they curated. These activities introduce learners to the importance of representative datasets.

Learning

AI learns from large sets of data to find patterns. As students build their own datasets in Modules 2-5, they are guided to evaluate the quality of the datasets they build. This includes exploring how different models function with different amounts of data and making connections between the size and quality of datasets and the quality of predictions and pattern recognition.

Natural Interaction

Many different types of data are necessary for AI to interact naturally with people. In Module 1, students are introduced to deep learning systems, like self-driving cars, that rely on multiple types of data to perform complex tasks necessary for natural interaction.

Societal Impact
AI can have positive and negative effects on society. Module 4 introduces students to the limitations of machine learning models, including how datasets can be biased, how that can affect predictions from AI systems, and what social ramifications this may have.

Method

This study employed qualitative and quantitative methods to explore the following research question:

- After participating in a computer vision curriculum, what are students' conceptions of AI, learning gains about AI, and changes in attitudes toward science?

Study context and participants
The teacher who is the focus of this work-in-progress paper is the 7th grade science teacher at a public school located in the southeastern United States affiliated with a large research university. She has 11 years of experience teaching science and has been at this school for 6 years. She taught the curriculum to one class (n = 27, 17 girls) at the beginning of Fall 2024. She chose to teach one module per day, in sequence and emphasized the portions of the curriculum, which aligned with the Nature of Science Next Generation Science Standards like *scientific knowledge is based on empirical evidence* and *science is a human endeavor* [14].

Curriculum description and modifications
The entire curriculum, including objectives, materials and guides, is available for free on our website [15]. Module 1, "Introduction to AI", was a general introduction to AI, and she heavily modified the original activities to only include a brief discussion about intelligence where students came to a shared definition of "logic and problem solving to adapt", and then a breakout activity where students independently learned about AI using resources curated by their teacher. These included watching a video, completing mini-lessons to explore how AI learns, and exploring how an AI tool called Class Companion used machine learning.

Module 2, "Shark Tooth Classification and Data Collection", was an introduction to shark teeth, paleontology, and data and databases. The teacher glossed over the heavy taxonomy and paleontology connections, and instead focused on the data and databases activities. She passed out shark teeth kits provided by the curriculum that included 15 real fossil shark teeth from 5 different species and one 3D printed Megalodon tooth. Students quickly sorted these into different categories according to functional morphology using a visual guide. She then guided students to think about different types of data and databases showing them two different databases that organize shark teeth: a crowdsourced database and a university database. Students explored the databases on their own chromebooks and through discussion came to the shared understanding that while both databases are important, they serve different purposes. Students were guided to recognize that the crowdsourced database had less data and fewer quality images, whereas the university database had more information regarding fossil location, measurement, and had higher quality images with less visual information in the background.

Module 3, “Training and Evaluating Machine Learning (ML) Models”, was an introduction to Google Teachable Machine (GTM) and instead of sorting shark teeth by hand, students built computer vision models using GTM designed to classify shark teeth. The module as written has students create two models, but the teacher had students create one, and really spend time on exploring and critiquing the data used for the model and the accuracy of the models. She connected this activity with the last class by having students discuss which database would yield the most accurate model. As students tested their models, she had them calculate the accuracy of each class and hypothesize why some classes had higher accuracy than others.

Module 4, “Identifying Biases and Limitations in ML Datasets”, was a deeper dive into how and why models can be biased. While the lesson guides give teachers pre-made biased models students can use, explore, and discuss, the teacher chose to show these models to the entire class and guide a discussion on how the data was curated and what bias the resulting model showed. She then had students evaluate their own models by showing them how GTM can create a confusion matrix. Again, students explored which classes were least accurate, brainstormed why, and then added or deleted pictures from their datasets to make them more representative.

In the final module, Module 5, “Conceive your own ML Model”, students were tasked with creating their own computer vision model that could correctly classify 80% of the time. Some of the models students created were those that classified leaves, human emotions, and colors.

Measures and data sources

Students’ conceptions of AI were captured using descriptive field notes the first author took while she observed the lessons and through reflection questions students answered on the last day of the module. Descriptive field notes and reflection questions were analyzed using thematic analysis [16]. Learning about AI was captured through pre-post assessments for each module created by the researchers and curriculum designers (Appendix A). Student attitudes towards science were captured using the reliable and validated Test of Science Related Attitudes (TOSRA) [17]. For this study, we explored four constructs on this scale, which included attitudes towards science, social implications of science, normality of science, and leisure interest. Both the learning assessments and TOSRA were analyzed using Wilcoxon signed-rank tests, because the assumptions for parametric tests were not met, which is common in survey research and with continuous variables in the context of a limited sample ($n = 27$) [18].

Findings and discussion

In reflecting on their conceptions about AI after the modules, all of the students described AI in a positive light and explained that AI could be “helpful.” Some students described how AI can be used to “solve problems”, but most students specified that AI can help with certain tasks like “identify fossils”, “identify different diseases”, and even help them “with their homework.” While all students had generally positive conceptions of AI, a majority ($n=15$) of the students acknowledged that AI was “biased” or had “limitations”, indicating that these students developed an emerging understanding of the social impact of AI [15]. These biases or limitations were often attributed to the fact that humans built the models or the models were created with incomplete and faulty datasets. When creating their own computer vision models, students recognized that “clear images” and “many images” were needed for accurate models. Many students recognized that AI models were only as good as the data used to build the model, with one student saying AI may have “trouble generating answers for people. If you give limited answers or bad pictures of descriptions.” Another student said “[AI] can have limitations because it only knows what we teach it.” These findings show students developed an understanding of how AI learns and reasons from representative datasets [6]. The teacher focused on the portions of the curriculum, which aligned with the Nature of Science standards, so that is potentially why students so readily recognized that computer vision models need representative and quality data.

These findings converge with the student learning data, as the students in the class showed significant learning gains for all the Modules except Module 1 (Table I). This is a promising result, which indicates that computer vision and AI can be taught to middle school students in formal, science classrooms. Module 1 was an introduction to AI, and at this school, students had already been introduced to AI in other classes through a school wide initiative. As indicated by the high means in the pre-test, it is possible that students already had a baseline understanding of AI. This, of course, has implications for the generalizability of these findings, and we are looking forward to exploring and comparing these results to other contexts.

The largest and most significant improvement was in Module 4, which emphasizes identifying bias and limitations in machine learning datasets and aligns with the “AI Big Ideas” about *societal importance* and *representation and reasoning* [6]. This converges with the student reflection data as over half of the students acknowledged computer vision models had limitations and could be biased. It is possible that these large learning gains may be attributed to the teacher’s focus on data and data curation, and we are looking forward to comparing these results with other teachers in other schools. Regardless, this finding is encouraging because educational scholars are concerned about the potential for AI in education to widen the digital divide and reinforce existing inequities. Educating students about how these machines can be biased not only introduces young people to how these systems work, but also gives them important context to the limitations and potential societal risks of these systems [19], [20].

TABLE I

DESCRIPTIVE STATISTICS FOR THE LEARNING ASSESSMENT AND RESULTS OF THE WILCOXON SIGNED RANK ANALYSES COMPARING PRE AND POST SCORES

Module	*M*		*SD*		*Z*	*p*
	Pre	*Post*	*Pre*	*Post*		
Introduction to AI	3.40	3.64	0.87	0.76	-1.40	0.172
Shark Tooth Classification and Data Collection	2.67	3.58	1.11	0.64	-3.20	0.001**
Training and Evaluating ML	3.00	3.44	0.83	0.51	-2.25	0.023**
Identifying Biases and Limitations in ML Datasets	2.36	3.30	0.86	0.97	-3.72	<.001**

Conceive Your Own ML Model	2.26	3.10	1.13	1.07	-2.61	0.007**

Modules were scored from 0 to 4.

Interestingly, while students' attitudes towards science changed after the curriculum, they did not do so significantly (Table II). Student attitudes toward science is an important K-12 education outcome as these attitudes have implications for lifelong learning and decision making regarding science [3], [4], [5]. But in this particular case, the means for each construct were already quite high. Further, the teacher implemented the curriculum very quickly (one module per day) at the beginning of the year, and this rapid implementation may not produce the magnitude of effect needed for significant changes in student attitudes. As we continue to observe more implementations (some of which are extended over a period of weeks and months) we will be examining differences in attitude development.

TABLE II

DESCRIPTIVE STATISTICS FOR THE TOSRA AND RESULTS OF THE WILCOXON SIGNED RANK ANALYSIS COMPARING PRE AND POST SCORES

Constructs	*M*		*SD*		*Z*	*p*
	Pre	*Post*	*Pre*	*Post*		
Attitudes Towards Science	3.26	3.19	0.55	0.72	0.30	0.779
Social Implications of Science	3.04	3.03	0.28	0.21	1.14	0.267
Normality of Scientists	3.09	3.19	0.35	0.28	-1.23	0.222
Leisure Interest	2.94	2.96	0.28	0.32	0.02	1.000

Likert scale questions were scored from 1 to 5.

Implications and future work

While curricular interventions using chatbots and LLMs have a growing literature base in education, curricula for young people on building and using computer vision models is sparse. In this paper, we explored the implementation of one 7th grade science teacher. While we found that students can learn about AI in formal settings, and particularly reflect deeply on the issues of bias and representative data, this is but one example of implementation. As this investigation is ongoing, we expect to collect more observation, reflection, learning, and student attitude towards science data from other classrooms in other schools. We expect this will provide a more varied and nuanced picture of the curriculum and implementations and its efficacy regarding important learning outcomes.

Appendix A

The following table details the learning assessment for each module and the connection to the "5 Big Ideas in AI". Each module has two constructed response and one open ended question, graded by the teacher.

Module	Items	"5 Big Ideas in AI"
1	1. Which of these traits is related to intelligence? a. Problem-solving b. Learning c. Memory d. All of the above 2. Which of these is an application of AI a. Spreadsheet b. Ruler c. Alexa or Siri d. Alarm clock 3. How do you use AI in your life?	3- Learning 1- Perception
2	1. Shark tooth shape relates to: a. Evolutionary relationships (Taxonomy) b. Tooth functions (Functional morphology) c. Diet (Ecology) d. All of the above 2. Why do scientists use databases? a. Find data b. Store data c. Share data d. All of the above 3. Why do paleontologists use tooth shape to classify fossils?	3- Learning

3	1. How do you train a machine learning model? a. Give it exercise b. Input data organized into classes c. Analyze model output data d. Write an equation 2. How do scientists organize data into different classes? a. They rank the data quality b. Depends on the research question c. Based on the size of the data d. Randomized groups 3. Why is it important for scientists to collect and prepare data before designing a machine learning model?	3- Learning 2- Representation & reasoning
4	1. Which factor might bias a machine learning model? a. Uneven sample sizes b. Human error c. Poor data quality d. All of the above 2. How do scientists evaluate a machine learning models’ accuracy? a. Increase model’s input data in the training dataset b. Remove biased input data from the training dataset c. Validate output results with a test dataset d. All of the above 3. How do scientists reduce bias in a machine learning model	5- Social impact 3- Learning 2- Representation & reasoning

5	1. When conceiving your own machine learning model to address a societal problem, what should be your primary focus during the planning phase? a. Determining the cost of implementation b. Defining the specific problem and identifying the data required c. Designing the user interface d. Choosing the color scheme for application 2. What is an essential aspect to consider when defining the classes for your machine learning model? a. The popularity of the classes among users b. The relevance of the classes to the problem being solved c. The ease of naming the classes d. The number of classes, keeping it minimal 3. What are some scientific problems that can be solved using a machine learning model? Why is it important to use machine learning to create this model?	5- Social impact 2- Representation and reasoning

References

[1] V.D., Păvăloaia and S.C., Necula. "Artificial intelligence as a disruptive technology—a systematic literature review" *Electronics* vol. 12, no. 5, 1102, 2023.

[2] Committee on STEM Education. "Charting a course for success: America's strategy for STEM education. National Science and Technology Council". 2018.

[3] R., Bybee and B., McCrae. "Scientific literacy and student attitudes: Perspectives from PISA2006 Science". *International Journal of Science Education* vol. 33, no. 1, pp. 7–26. 2011. doi:10.1080/09500693.2010.518644.

[4] E.W., Jenkins and R.G., Pell. "The relevance of science education project (ROSE) in England: Asummary of findings. Leeds: Centre for Studies in Science and Mathematics Education". *International Journal of Science Education.* 2006.

[5] Organization for Economic Cooperation and Development. "Informe PISA 2006,Competencias científicas para el mundo del mañana". [PISA Report 2006, science competencies for tomorrow's world]. Madrid: Santillana.

[6] AI4K12. "AI4K12*"*. https://ai4k12.org (accessed Jan. 10, 2025).

[7] M., Gatlin. "Assessing Pre-service Teachers' Attitudes and Perceptions of Using Artificial Intelligence in the Classroom". *Texas Educator Preparation* vol. 7, no. 2, pp. 1-8. 2023.

[8] J., Nayak, T., Keane, T., Linden & A., Molnar. "Teaching high school students artificial intelligence by programming Chatbots". *Teaching Coding in K-12 Schools: Research and Application* Cham: Springer International Publishing. 2023, pp. 263-276.

[9] Y., Song, X., Tian, N., Regatti, G.A., Katuka, K.E., Boyer & M., Israel. *Artificial Intelligence Unplugged: Designing Unplugged Activities for a Conversational AI Summer Camp.* In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1. March, 2025, pp. 1272-1278.

[10] M., Yue, M.S.Y., Jong, and Y., Dai. "Pedagogical design of K-12 artificial intelligence education: A systematic review." *Sustainability*, vol. 14, no. 23. 2022.

[11] Moore, R. L., Jiang, S., & Abramowitz, B. (2023). What would the matrix do?: A systematic review of K-12 AI learning contexts and learner-interface interactions. Journal of research on Technology in educaTion, 55(1), 7-20.

[12] AI4K12. "Grand Band Progression Charts". https://ai4k12.org/gradeband-progression-charts/ (accessed Jan. 10, 2025).

[13] AI4K12. "Five Big Ideas in AI Figure." https://ai4k12.org/gradeband-progression-charts/ (accessed Jan. 10, 2025).

[14] Next Generation Science Standards. "Middle School Storyline PDF." Retrieved from https://www.nextgenscience.org/overview-topics (accessed Jan. 10, 2025).

[15] https://education.ufl.edu/sharkai/

[16] V., Clarke and V., Braun. “Thematic analysis”. *The Journal of Positive Psychology,* vol. 12, no. 3, pp. 297-298. 2017.

[17] B.J., Fraser. “Test of science-related attitudes”. *Science Education*. 1981.

[18] M., Kraska-Miller. “Nonparametric statistics for social and behavioral sciences”. CRC Press. 2013.

[19] K., Holstein and S., Doroudi. “Equity and artificial intelligence in education: will" aied" amplify or alleviate inequities in education?” 2021. arXiv preprint arXiv:2104.12920.

[20] M., Roshanaei, H., Olivares & R.R., Lopez. “Harnessing AI to foster equity in education: Opportunities, challenges, and emerging strategies” *Journal of Intelligent Learning Systems and Applications* vol.15, no. 04, p.p. 123-143. 2023.