Received: 10 April 2025 | Revised: 21 May 2025 | Accepted: 27 June 2025

DOI: 10.1002/aaai.70013

SPECIAL TOPIC ARTICLE

Artificial intelligence education in Georgia middle schools

David S. Touretzky[1] | **Christina Gardner-McCune**[2] | **Bryan Cox**[3] | **Judith Uchidiuno**[3] | **Xueru Yu**[3] | **William Gelder**[3] | **Tom McKlin**[4] | **Taneisha Lee Brown**[4] | **Bejanae Kareem**[5] | **Woojin Chung**[6] | **Amber Jones**[7] | **Janet Kolodner**[8]

[1]Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

[2]Department of Computer and Information Science and Engineering, University of Florida, Gainesville, Florida, USA

[3]College of Computing, Georgia Institute of Technology, Atlanta, Georgia, USA

[4]The Findings Group, Decatur, Georgia, USA

[5]BK International Education Consultancy, Atlanta, Georgia, USA

[6]Ewha Womans University, Seodaemun-gu, Seoul, South Korea

[7]Amber Sparks Education LLC, Lithonia, Georgia, USA

[8]Carolyn A and Peter S Lynch School of Education and Human Development, Boston College, Chestnut Hill, Massachusetts, USA

Correspondence
David S. Touretzky, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA.
Email: dst@cs.cmu.edu

Funding information
National Science Foundation awards, Grant/Award Numbers: DRL-2049029, DRL-2048502; Google

Abstract

In a partnership between four universities, the Georgia Department of Education, and nine Georgia school districts, we developed a 9-week middle school elective called “Living and Working with Artificial Intelligence,” and a professional development (PD) program for prospective middle school AI teachers. To ensure that our curriculum could meet the needs of all learners, we recruited a diverse set of districts that included rural districts serving mainly White students, urban districts that were majority African American, and suburban districts serving a mix of Hispanic and African American students. Now in its fourth year, our “AI for Georgia” project (AI4GA) has provided PD to 20 teachers and AI education to over 1600 students. The AI4GA curriculum does more than foster AI literacy: It empowers students to view themselves as creators of AI-powered technology and to think about future career options that involve the use of AI. The project is now expanding to schools in Texas and Florida. In this article, we review the history of the project, discuss our co-design process with our teachers, and present results from studies of teacher PD and student learning.

INTRODUCTION

As the impact of artificial intelligence on daily life continues to grow, the need for early AI education is increasingly apparent. A recent White House Executive Order on K-12 AI education explains that fostering AI competency in our youth is essential for ensuring that the United States remains a global leader in the field (Trump 2025). Some recent AI education efforts target high school students, but the AI4K12 Guidelines begin in kindergarten, and it is worth noting that by the time children arrive in kindergarten, many have spent 2 years conversing with Alexa or similar agents. We decided to tackle the problem of AI education in middle school, meaning grades 6–8, as it provides an opportunity to introduce students to AI in ways that we hope shape their interest in potential AI-related careers. Middle school is often touted as the best time to shape students' career interests and to reduce declining interest and achievement in STEM fields (Almeda and Baker 2020; Trotman 2017). To do that well, it is important to actively include teachers in the design of the curriculum, to ensure it addresses the needs of all middle school students. This article tells the story of what we accomplished and what we have learned so far.

AI4K12 creation and work

The AI4K12 Initiative (AI4K12.org) launched in 2018 as a joint project of the Association for the Advancement of Artificial Intelligence (AAAI) and the Computer Science Teachers Association (CSTA), with funding from the National Science Foundation. The project was initially focused on addressing a gap in the CSTA K-12 Computer Science Standards (Computer Science Teachers Association 2017) that are implemented or adapted by states across the United States. The AI4K12 Initiative developed national guidelines for teaching AI in K-12 and published a list of "Five Big Ideas in AI" (Touretzky, Gardner-McCune, Martin et al. 2019) that has been widely adopted by K-12 educators and researchers both nationally and internationally. An infographic poster explaining the five big ideas is available at AI4K12.org and has been translated into 16 languages. In 2019, the AI4K12 team was approached by the State of Georgia and districts in Florida to consider how to shape K-12 AI education at the state level.

Georgia's state AI initiative

In 2019, the state of Georgia approved K-8 standards for Computer Science (CS), expanding opportunities for students to learn CS prior to high school and transitioning the subject into a K through 12 discipline. Later that year, a state education CS advisory council met to discuss the evolution of K-12 CS in Georgia. On the docket were Artificial Intelligence, Data Science, and E-Sports. From this meeting, the state created a high school Career, Technical, and Agricultural Education (CTAE) pathway for AI, the first of its kind. At the time, the council decided to postpone the creation of middle school AI standards until there was more clarity on what AI education would look like in middle school. Around that time, the state CS program lead was approached by the AI4K12 team to collaborate on a proposal to develop an AI curriculum for middle school. This provided a natural opportunity to explore the needs for AI education at the middle school level and to provide a testbed for translating the AI4K12 guidelines into curriculum.

State symposium for AI

Georgia was one of 27 states and three US districts/territories to participate in AI4K12.org's 2-day January 2021 State of K-12 AI Education in Your State Workshop. The Georgia state CS program lead at the time, Bryan Cox, was part of the organizing committee. The purpose of the workshop was to guide state delegations to develop a shared vision of why it is important to incorporate AI education into their state's CS and broadening participation in computing plans. Each state delegation formulated an action plan to achieve the goals they set for their state. The workshop was co-facilitated by Leigh Ann Delyser, CSforAll co-founder and former executive director, based on an AI-adapted SCRIPT District Training program (CSforALL 2024). As part of the Georgia state plan, there was a focus on developing a progression of AI education as a strand throughout K-12 CS instruction. There was also an intent to integrate AI instruction across other content areas. All of these pieces laid the groundwork for the AI4GA project.

PROJECT OVERVIEW

AI for Georgia

The AI4GA ("AI for Georgia") project (Figure 1) is a 4 year, $1.5 million NSF ITEST (Innovative Technology Experiences for Students and Teachers) project that began in 2021 to develop and pilot a 9-week AI elective for Georgia middle school students called "Living and Working with Artificial Intelligence." To accomplish this, we also had to conduct in-person professional development (PD) training for Georgia teachers. To date, the project has provided PD for 20 middle school teachers and AI instruction to over 1600 students in 15 schools in nine districts across the state. Five of the teachers have gone on to become teacher leaders who are now mentoring other teachers interested in teaching AI. The project is now expanding to include schools in Texas and Florida.

FIGURE 1 Artificial Intelligence for Georgia (AI4GA.org) logo, with the state drawn in peach color (Georgia is "The Peach State") and the partial outline of a cyber/human profile mirroring the AI4K12.org logo.

Research practice partnership structure

While AI4GA was funded under the ITEST program, the partnership structure aligns well with NSF's Research Practice Partnership (RPP) framework. RPPs are a currently recommended structure for NSF projects that leverage multiple sources of expertise relevant to a given educational context (Henrick et al. 2017; McGill et al. 2021). The researchers bring experience in developing projects that collect data and construct knowledge, while the practitioners bring expertise with the educational setting and the nuances of operating therein.

The principal investigators on the AI4GA project are two university CS professors (from Carnegie Mellon and the University of Florida) and the CS lead for the Georgia Department of Education. While the researchers are able to draw on the academic literature and bring their own deep knowledge of AI education to bear, the state education lead offers insights on sustainability strategies within the K-12 education system, recruitment and support strategies for teachers, and a variety of professional learning modalities common in K-12 education. Together, the PI team offers both content and context expertise.

Other researchers on the project provided expertise on co-design strategies (discussed further below), culturally relevant pedagogy, project evaluation methodology, and project management. Each of these were critical project components. The team's ability to rely on each other to lead in their areas of expertise was a significant factor in the project's success.

Finally, the middle school teachers were essential partners. While the state CS lead provided a big picture contextual perspective, it was the middle school teachers who helped translate the content knowledge of the researchers into a viable curriculum for middle school teachers and students. The co-design process, including extensive and ongoing professional learning experiences, took teachers from knowing very little about AI concepts to confident advocates for AI tools, demonstrations, and activities that would be effective with middle school students. The teachers and their students were drawn from a variety of communities, so teachers in rural areas found content (e.g., real-world AI applications or interesting videos) that connected best with their students, while teachers in suburban or urban school systems did likewise.

The district partners were located throughout the state of Georgia, providing insights from rural school systems (Thomas County, Tift County, Burke County), suburban school systems (Clayton County, Douglas County, Fayette County), and urban school systems (Muscogee County, Dougherty County, and Atlanta Public Schools). This made for a rich project team, and products that have a broad impact and utility.

TEACHER PROFESSIONAL DEVELOPMENT

Middle school CS teachers typically come to the project with no knowledge about artificial intelligence. Thus, at the start of the project, we conducted PD activities to familiarize our first teacher cohort with the basics of AI before we could engage them as curriculum co-designers. This is not unlike other teacher PD models developed for introducing non-CS teachers to CS (Goode, Margolis, and Chapman 2014; Price et al. 2016; Rosato et al. 2017). Neither the PD nor the co-design activities were one-shot events. Both are ongoing processes that continue to this day, not just for new teachers recruited to the project, but also for our continuing teachers. Both activities have evolved as the project matured. In the following sections, we first describe the trajectory for the PD component and then the trajectory for co-design, but we want to emphasize that after initial training, these become overlapping activities.

Introductory workshop

Our initial PD model was a week-long workshop during the summer, followed by weekly check-ins throughout the school year. In the first year, the introductory workshop provided a broad survey of AI, focusing on topics relevant to the planned curriculum and the many ways AI technologies contribute to daily life. During this time, an initial draft of the curriculum was being written by the

PIs to provide a starting point for the co-design process. In subsequent years, the week-long introductory workshop used the actual curriculum materials, providing teachers the opportunity to experience the course as a student would. These later workshops were co-taught by the PIs and current teachers, so new recruits could ask questions of teachers who had already taught the curriculum about how to adapt it for their students. We also now provide time in the PD for teachers to plan out their implementation. In this way, when teachers leave the PD, they have a plan for how they will implement the curriculum in their classroom. They also spend some time adapting one of the lessons within the workshop so they can get feedback on their work. They leave the workshop with contact information for all the other teachers in the cohort, as well as the PIs and teacher leaders.

Implementation support

The biggest risks after a 1-week intensive PD are that teachers will not go on to teach the curriculum, or will struggle to implement it, or will not have the support they need to get timely answers to questions. To address these issues, we provided intensive implementation support during the school year in the form of weekly check-ins via Zoom. This provides opportunities for teachers to talk about their lessons right before implementation and then immediately after implementation, drawing on experiences of other teachers who may have also just implemented the lesson. It also provides valuable feedback to the research team. In addition, we used some of this meeting time to provide refreshers on content from the 1-week PD, especially how to use the many AI tools and demos the curriculum draws upon. In the fourth year of the project, when we did not have a new cohort, we met with our teachers every other week since they did not require as much support. With our first cohort of Texas teachers, we are using a different model consisting of a 4–6-h Saturday session once per quarter, in accord with what WeTeach_CS has been using for its professional learning community.

From teachers to teacher leaders

Teacher leaders offer professional learning opportunities to other teachers and act as mentors or coaches. Developing a corps of AI teacher leaders is key to our strategy for scaling up the AI for Georgia project and ensuring long-term sustainability. This has been a successful approach to scaling CS PD nationally with projects such as Exploring Computer Science, Mobile CS Principles, and Beauty and Joy of Computing (Goode, Margolis, and Chapman 2014; Price et al. 2016; Rosato et al. 2017).

We are currently on our second iteration of the teacher leader PD model. Initially, teacher leader PD focused on co-leading the summer introductory workshop for new teachers. However, in Summer 2024, we received a grant from Google that allowed us to explore a more comprehensive approach. We held a week-long workshop for five of our Georgia teacher leaders to help them better understand the depth and scope of AI content and pedagogical knowledge needed for training new teachers. The workshop focused on the following: deepening teachers' understanding of AI through reviewing AI concepts, AI4GA curriculum modules and assignments, learning objectives, and student work; recognizing common misconceptions and stages of student learning of AI concepts; sharing strategies and best practices for implementing specific activities and for addressing student and teacher misconceptions; and improving the design of assessments of students' AI knowledge and skills.

The activities in the workshop followed a modified Student Work Analysis Protocol to discuss and analyze student work (see Rhode Island Department of Education & National Center for the Improvement of Educational Assessment n.d.). The workshop outcomes included refined activities, assessments, and rubrics that deepened both the students' and the teacher leaders' understanding of AI. In refining these materials, the teachers were also continuing to contribute as co-designers on the project.

Teacher leader practicum

After our first cohort of teachers completed their first year of teaching, we began transitioning the facilitation of the summer teacher PD to the teachers, where they assumed the roles of teacher leaders and mentors. Seeing the teachers as the best source of knowledge about how to teach the curriculum to middle school students, the researchers shifted into a content knowledge advisory role, providing just-in-time explanations of AI or how to use AI tools. The first summer where the teacher leaders assisted in the PD facilitation, we had a 50/50 split between researchers and teacher leaders leading the professional learning. The second summer, we moved to a 20/80 researcher/teacher split, with the researchers focused on training on new tools and curriculum materials that had been developed to address project needs.

As we continue refining our teacher leader development program, we see several benefits of the teacher leaders assuming more responsibility for training new AI teachers. This practicum component helps teacher leaders to: increase their knowledge of the curriculum and become more confident in their AI knowledge; better understand the needs of novice AI teachers during initial PD; make explicit their pedagogical practices for teaching AI and engaging their students; and create their own resources to use when facilitating PD sessions, either within the AI4GA summer introductory workshops or in other venues such as conferences or state-sponsored education events.

PD led by teacher leaders helps teachers to: see the curriculum modeled by peers, which improves their learning and retention; increase their comfort asking questions about content knowledge and pedagogical strategies; gain clearer understanding about what practical and instructional needs and barriers they may encounter in the classroom; increase their confidence by knowing that the materials have been implemented and refined by teacher leaders; and feel more comfortable making adaptations for their teaching style and student needs.

CO-DESIGNING THE CURRICULUM

Co-design is a collaborative creative process that purposefully includes diverse stakeholders to co-create a product or innovation addressing a shared need. By actively including multiple perspectives (Roschelle and Penuel 2006), and by establishing a sense of equal power and open communication among participants, innovation and ideas become stronger, more nuanced, and more robust. In the case of AI4GA, our intent was to co-design with teachers an AI curriculum that could effectively meet the needs of diverse middle school learners across a varied range of learning environments in Georgia, each with students at varying levels of academic proficiency and CS literacy. The sequence of AI topics covered in the curriculum will be discussed in a later section, "Curriculum Content."

The co-design process with our first cohort of teachers was lengthy, since we were starting from scratch. The participants were five university researchers, three curriculum and PD specialists, two evaluators, and five middle school teachers. Over the course of 33 weeks, the team met weekly for 1-h sessions. The process took place in the following three phases, detailed below: Phase 1: ideation and framework development; phase 2: piloting and adaptation; and phase 3: refinement and expansion.

Phase 1: Ideation and framework development

This initial co-design phase, lasting 12 weeks in the Fall of 2021, began with an examination of some draft materials put together by the principal investigators. Researchers posed the following practical questions: Do the concepts make sense? Would this work in your classroom? What resources do you need to teach this content effectively? How can we make this curriculum accessible to teachers who haven't participated in professional development?

Teachers found the initial materials unsuitable for middle school due to heavy reliance on a lecture slide format and insufficient offerings of student activities as vehicles for learning. What had worked for educating adult teachers would not work for their 11–14-year-old students. This critical feedback from the teachers triggered discussions and brainstorming sessions that reshaped the curriculum's design. As the researchers prepared later curriculum modules for discussion, they took these lessons to heart and worked to limit the number of slides and propose more activities students could engage in, either online or with paper and pencil. The teachers continued to critique these materials and suggest changes to improve their effectiveness.

Another point the teachers emphasized was the importance of giving students choices when designing activities. For example, they might be offered a collection of robot videos and allowed to pick three for the class to analyze. Or they might be offered multiple ways of expressing what they have learned about self-driving cars by writing a paragraph, making a drawing, or creating a comic strip (using an online tool).

Phase 2: Piloting and adaptation

In the Spring of 2022, we entered the second co-design phase, lasting 16 weeks, in which teachers held their first pilot offerings of their Living and Working with Artificial Intelligence course. Although all of these were 9-week AI elective courses (called "Connections" courses in Georgia), the format varied. In some schools, these classes met on alternate days for longer periods, while in others, the classes met daily for shorter periods. Teachers also had different teaching styles. Some teachers have their students work mostly independently, with the teacher supervising, while others place more emphasis on group activities. Also, some teachers needed to accommodate students with low reading levels, or who were second language learners with limited English proficiency.

During this phase, key questions included the following: How are teachers adapting the materials? What components of the curriculum engage students? What challenges are teachers encountering? and What additional resources are needed?

Each teacher adapted the materials to best suit their teaching styles and the needs of their students. They created more student-friendly versions of the lesson slides from the PD materials by adding visuals and styling to improve accessibility and engagement. The student-facing decks were also considerably shorter than the teacher decks the researchers provided. In general, teachers adjusted their lesson pacing based on what they estimated could be accomplished in a class period and developed supplemental materials such as worksheets and unplugged activities (activities that do not require technology).

One of the first cohort teachers working in a rural school adapted the lessons to better suit the background knowledge and interests of his students. For example, in the Unit 1 activities on autonomous robots and self-driving vehicles, he included agriculture-related robot content, such as self-driving tractors. Furthermore, to engage the many advanced learners in his class, he regularly structured his lessons with brief introductions to the activity, before giving students more independence to explore the content and work on challenges while he maneuvered about the room, engaging students with questions and guidance.

Another teacher from the same cohort created many worksheets to function as guided notes and activities. Her students consisted of many English language learners and students with low reading proficiency. To accommodate, she added slides to presentations that included language translations of vocabulary or other key information, found videos and materials related to content in other languages, and generally structured her class with whole-group readings and discussion of the background content. Then activities would often be completed as partnered activities or with guided practice before students attempted the content independently.

A third teacher from the first cohort regularly emphasized students' interests and career aspirations in her adaptations. In her classroom, students took surveys at the beginning of the year to discover their learning styles and career interests. Her adaptations of the lessons then offered students choice in how they engaged the content, such as articles or videos, and how they demonstrated their learning, such as making an animation or doing a presentation. She encouraged her students to engage with the content in the modality that best suited their learning style and interests based on the survey data. She also structured her lessons to ensure that students connected AI content to careers and their interests.

The teachers shared these materials with the research team and each other during the weekly implementation support meetings. During this phase, researchers shifted to being observers and coaches, monitoring how the curriculum was implemented and gathering feedback on its strengths and weaknesses. The teachers, taking a more active role in these meetings, would explain how they had engaged the students, student reception and misconceptions, and any materials or activities they had created. Because the teachers taught the content at different paces, often the materials and activities of a teacher further along in the curriculum would be borrowed and further adapted by another teacher when they finally arrived at that content. Because the middle school courses last 9 weeks, these weekly meetings allowed teachers to learn from their own experiences and the experiences of others in their cohort, receive feedback, and then adapt and improve the materials before the next 9-week term. This process during the Spring of 2022 created many of the foundational materials that the current refined materials are based on.

Phase 3: Refinement and expansion

The final phase, lasting 5 weeks during the Summer of 2022, focused on reviewing and refining the curriculum based on insights gained during the pilot phase. Teachers took the lead in proposing new ideas and making adaptations to improve the curriculum. Key goals included re-scoping AI concepts, expanding the number of student activities, creating lesson plans for teachers new to the curriculum, and addressing challenges related to student engagement and content relevance.

During this phase, researchers and teachers asked the following: What is the big picture for this curriculum, unit, module? What are the foundational AI concepts students should learn? How do we scope the content for middle school students? Is this the right learning objective? What is the appropriate pacing: 1, 2, 3, or 4 days for a module? What is the connectivity between modules within a unit? How and when do we revisit concepts? For each activity, do we keep, refine, or remove it?

During this phase, teachers were positioned as experts in both teaching and curriculum development. The team reviewed each unit in detail, identifying areas for improvement and brainstorming solutions. This time also provided an opportunity for teachers to ask questions about AI content they were still unsure of. By the end of this phase, the curriculum had been fine-tuned to better meet classroom needs and implement teacher feedback. Teachers felt more comfortable and confident about the curriculum and their AI knowledge.

Co-design as iterative refinement

While the first year of co-design was particularly lengthy and intense, we maintained a modified co-design model in subsequent years, focusing on phases 2 and 3. New teachers who offer the course for the first time are encouraged to adapt it to their needs, with intensive implementation support through weekly meetings. In this way, they become co-designers as well as implementers. Then, in the summer (phase 3), we take stock of what has been learned and make further improvements to the course, which can include adopting teacher-created materials into the official reference version of the curriculum hosted on the project website.

We conclude this section with two examples of significant course innovations that could only have been achieved through co-design (Gelder et al. 2025). The first is our treatment of word embeddings. Word embeddings are high-dimensional vectors used to encode word meanings in large language models. One of the researchers developed an online tool for interactive visualization of word embeddings that was suitable for K-12 use (Bandyopadhyay et al. 2022), and this tool was the initial focus of the lesson plan. To help students understand the notion of a semantic feature space, the lesson begins by introducing simplified 2D and 3D feature spaces that can be displayed as conventional graphs. But this is still fairly abstract for middle school students. One of the teachers devised an unplugged activity where they used the physical classroom to represent a 3D feature space, with the first dimension aligned with the north and south walls, the second with the east and west walls, and the third with the floor and ceiling. Given a word, students individually position themselves to indicate where that word should be located in the space, and compare their results. Another teacher then made refinements to this activity, and other teachers quickly adopted it. The exercise has now been incorporated into the master curriculum and is taught to all new teachers. Once students have physically navigated a 3D semantic feature space, they are better equipped to consider more abstract representations, such as a 3D graph or the 300-dimensional embedding used in the online demo.

The lesson on word embeddings also includes playing Semantris, an engaging browser-based game based on semantic similarity (Google Research 2018). Students are asked how they think Semantris knows that words like "pasta" and "spaghetti" are related. They learn that Semantris uses word embeddings to measure similarity.

The second co-design example is the My Dream Bot exercise. This began as a suggestion from one of the researchers during a brainstorming session that students should be invited to design their "dream robot" as a way of demonstrating their understanding of concepts such as robot sensors and automated decision making. The team devised a series of activities students could engage in over the course of several weeks to select a problem that could be solved by an autonomous robot and work out details such as the sensors the robot would use, how the robot would use route finding and computer vision to navigate the environment, and the societal impacts of the robot. Due to time limitations as well as resources, students are often unable to build physical models of robots. Thus, one of the activities asks students to create a visual depiction of their robot. The co-design team explored multiple ways of accomplishing this, including freehand drawing or the use of a drag-and-drop collection of robot components. Some students had prior experience with a 3D modeling program (Tinkercad), so this was also included as an option. My Dream Bot gives students an opportunity for personal expression while demonstrating what they have learned about autonomous robots. It has also proven valuable as an assessment tool (Yu et al. 2025). Most students were able to explain that camera, LIDAR, and RADAR would help the robot see things. However, although their robots were designed with distinct functions and operated in specific environments, students did not specify what things the sensors looked for (e.g., traffic signs, pedestrians, buildings). Students also leveraged their prior computing experiences in robotics clubs to incorporate sensors not covered in our curriculum, such as distance and pressure sensors. Overall, the Dream Bot activity provides a summative assessment of students' knowledge of how autonomous robots make decisions using data from their sensors.

CURRICULUM CONTENT

The AI4GA curriculum is aligned with the "Five Big Ideas in AI" developed by AI4K12.org (Touretzky, Gardner-McCune, Breazeal et al. 2019;Touretzky, Gardner-McCune, Martin et al. 2019). Students are introduced to these big ideas in the first lesson. They are summarized in the infographic shown in Figure 2. But we do not proceed through them sequentially. Each course unit touches on multiple big ideas, with societal impact and ethical design (Big Idea 5) always attended to. Our design is student-focused in that we choose topics relevant to students' lives and interests, and incorporate opportunities for students to express themselves and demonstrate competence through project-based learning.

The resulting curriculum allows students to see diverse faces in AI and positions students as AI users, decision makers, problem-solvers, and creators. It also encourages them to explore and debate the impacts of computing. Our curriculum centers on the fair and ethical consideration and treatment of all people and invites students and teachers to think about the wide range of impacts AI will have on individuals and different groups within our society.

The curriculum is organized as three units, each subdivided into multiple modules as shown in Tables 1–3. Unit 1 covers autonomous robots and self-driving vehicles. Unit

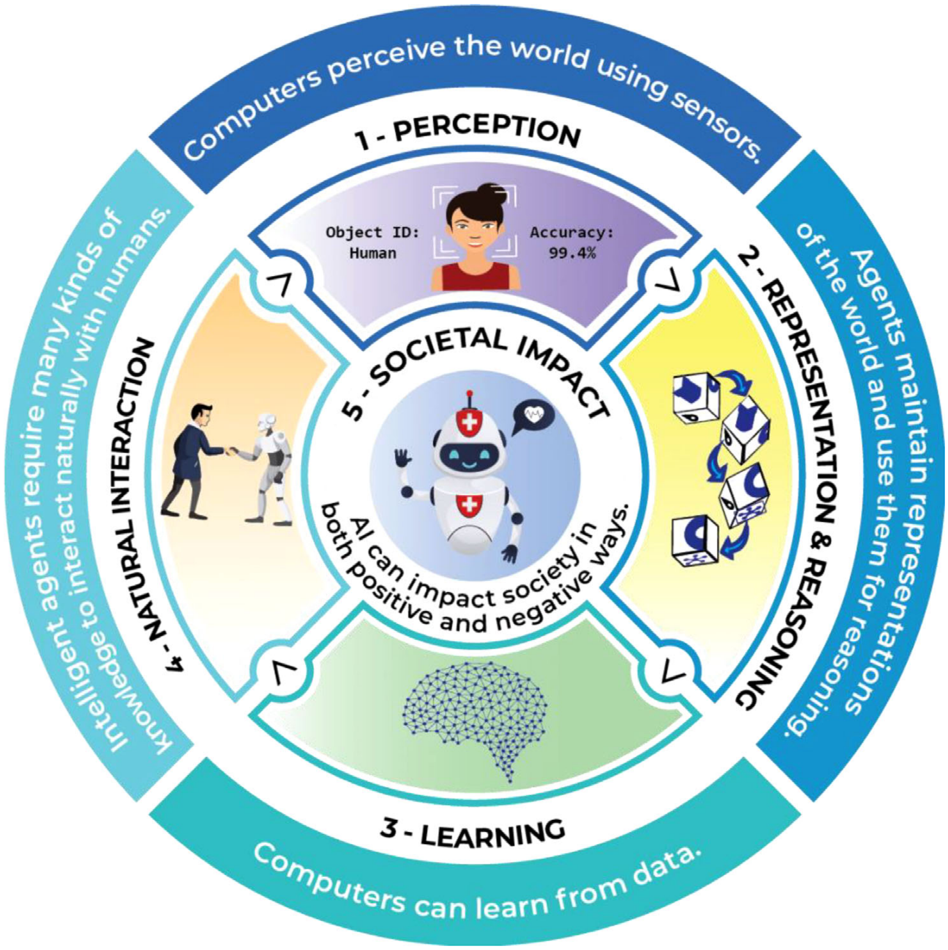

FIGURE 2 The Five Big Ideas in AI, courtesy of AI4K12.org.

TABLE 1 Modules from Unit 1.

Unit 1: Autonomous robots and self-driving vehicles	
Module	**Description**
1.1	**Course overview**; What is AI? The Five Big Ideas in AI
1.2	**Autonomous robots**: What is autonomy? Illustrative robot videos; “Robot or Not?” unplugged activity
1.3	**Anatomy of a self-driving car**: Sensors, subsystems, types of decision-making by autonomous vehicles
1.4	**Robot and computer perception**: Pixels and images; feature extraction; how face detection works; building a face filter in Scratch
1.5	**Route finding**: Graphs and trees; reasoning algorithms; breadth-first search
1.6	**Case study**: Sleeping drivers in self-driving cars
1.x	**End of unit mini project**: My Dream Bot (designing an autonomous robot)

2 investigates how computers understand language. Unit 3 covers machine learning and automated decision making.

Due to teacher and student pacing and other time limitations, it was not possible to cover all three units in 9 weeks. All classes included Unit 1. Then, in some semesters, teachers covered Unit 2, and in others, they covered Unit 3, but some teachers chose to teach selected modules from both units.

Fundamental understandings

In designing the curriculum, we wanted students to come away with fundamental understandings about AI that are more than superficial. We list three of them here.

Sensing versus perception (Big Idea 1). Perception is the nontrivial extraction of meaning from raw sensory signals. Students should appreciate the difference

TABLE 2 Modules from Unit 2.

Unit 2: How computers understand language	
Module	**Description**
2.1	**Unit overview**: Understanding language
2.2	**How do computers understand what we say?** Waveforms, spectrograms, speech recognition
2.3	**How do intelligent assistants understand and answer questions**? Parsing, syntax, semantics, search queries
2.4	**Word embeddings**: Semantic feature space; coordinates of a word
2.5	**How computers represent and generate meaning (optional)**: Machine translation, idioms, and colloquialisms
2.6	**Sentiment analysis**: Determining the emotional valence of text
2.7	**Chatbots**: Types of chatbots; creating your own chatbot
2.x	**End of unit case study: ChatGPT**: Examination of the benefits, limitations, and drawbacks of using ChatGPT, culminating with an in-class debate

TABLE 3 Modules from Unit 3.

Unit 3: Machine learning and automated decision making	
Module	**Description**
3.1	**How do computers make decisions?** Types of reasoning problems (classification, prediction, recommendation, planning and scheduling); types of reasoning algorithms; “Name the reasoner” activity
3.2	**How can a computer learn to classify objects from examples?** Decision trees; “Candy Land” classifier activity
3.3	**How do we train a computer to make decisions?** A fishy feature space; machine learning pipeline; training versus inference; automated decision tree learning with MachineLearningForKids
3.4	**Machine learning with datasets**: Training a classifier or predictor using Code.org’s AI Lab; classification versus prediction; training set versus test set
3.5	**Neural networks**: Weights, activations, thresholds, and layers; use of Neuron Sandbox to understand decision making by linear threshold units
3.6	**Case study: Does AI make better decisions than people?** Use of automated decision-making systems that affect people; identifying values of diverse stakeholders; bias, fairness, and transparency
3.7	**Machine learning mini project**: Use one of three tools (Teachable Machine, MachineLearningForKids, or AI Lab) to create a classifier or predictor using your own data, paying attention to ethical design considerations such as bias, fairness, and transparency

between a video camera recording images versus a computer vision system that recognizes objects in the world, or a microphone that senses sound versus a speech-based computer interface. They learn that computer perception is the product of sophisticated algorithms plus extensive domain knowledge. We explore this topic in discussions of self-driving cars, autonomous robots, and face recognition in Unit 1, and experiments with speech understanding systems in Unit 2.

Types of reasoning (Big Idea 2). Following the AI4K12 Guidelines, we define specific types of reasoning problems that AI is used for. The curriculum focuses on the following four: classification, prediction, recommendation, and planning and scheduling. Classification and prediction are similar; classification assigns inputs to discrete categories, while prediction, sometimes referred to as regression, estimates continuous values. Students experiment with classification in Unit 3 by constructing and hand-simulating decision trees to classify candy. They encounter it again when using Teachable Machine (Phillips 2019) to classify hand gestures. Students are already familiar with recommendation systems through apps such as Facebook and Tiktok, although they may not understand how they work. Without getting into technical details, we explain that recommendation systems select items that were liked by people who have liked other things that the user likes. Planning and scheduling are examples of *search*, a broad topic in AI that we only briefly touch upon due to time constraints. In Unit 1, students hand-simulate a breadth-first search algorithm to solve route-finding problems for a self-driving car.

We reinforce these concepts through a “name the reasoner” activity where students are given 15 examples of AI applications and asked what type of reasoning each one involves. The recent explosion of large language models with more general reasoning capabilities may lead us to refine our treatment of reasoning in the future.

Supervised learning (Big Idea 3). A crucial distinction we want students to make, often overlooked in introductory AI treatments, is the difference between training and inference. The job of a machine learning algorithm is to *construct* a reasoner, such as a decision tree or a neural network, during a training phase. It might do this by adding nodes to the decision tree or by adjusting weights in the neural network. The learning algorithm is not itself the reasoner. Once the reasoner has been constructed, it can be applied to novel inputs; this is the inference phase.

In Unit 3, students hand-simulate decision tree learning in a Candy Land activity inspired by Lee et al.'s PastaLand (Lee et al. 2021). When using Teachable Machine, they learn about the importance of collecting adequate training data and the effects of biased training data. Optional computer-based activities in Unit 3 include the use of MachineLearningForKids to construct a decision tree classifier, and the use of Code.org's AI Lab to construct a classifier or predictor using a large dataset and then measure its prediction accuracy.

Importance of mental models

Another of our curriculum design principles is that we should help students acquire *mental models* of important AI concepts. Although middle school students have limited mathematical experience, they are capable of engaging with sophisticated concepts when appropriately scaffolded. Examples from the AI4GA curriculum include tree and graph data structures, feature spaces, embeddings, and linear threshold neurons.

Tree and graph data structures are too sophisticated for middle school programming courses, but are commonly introduced in unplugged activities. Students encounter decision tree classifiers in Unit 3, and search trees in the Unit 1 route finding activity. Route finding also shows them how a road map can be represented as a graph composed of nodes and links. Students learn the vocabulary for describing these structures and see how they can be represented visually. They are key data structures for symbolic AI, supporting the representation part of Big Idea 2 (Representation and Reasoning).

Feature spaces and feature vectors. We use an unplugged activity, described earlier, to introduce students to the notion of a semantic feature space. Feature spaces are also referenced in the Candy Land classifier activity where students make up their own features, in an unplugged activity called Fishy Feature Space where individual fish are described by feature vectors (e.g., values for body shape, tail shape, color, etc.), and in an optional AI Lab machine learning exercise where each dataset has an associated feature space. Embeddings, discussed previously, are taught as a generalization of this concept where the number of dimensions (features) is very large and the meaning of each dimension is not necessarily known to us.

Neural networks is perhaps the most challenging topic for which we want students to develop a mental model. Showing them a picture of a multilayer perceptron gives no real insight, and telling them that neural networks "work like the brain" is misleading, since we do not actually know how the brain works. But the details of how a multilayer perceptron actually works are too complex and abstract for middle schoolers. Therefore, we decided to focus on how a single neuron can make simple decisions, such as deciding whether you can make a peanut butter and jelly sandwich, which requires having both peanut butter and jelly. We use a browser-based tool called Neuron Sandbox (Touretzky, Chen and Pawar 2024) to simulate the operation of a linear threshold unit, and developed a highly scaffolded approach to teach students to reason about how this unit computes (Touretzky et al. 2025). Students learn to solve reasoning problems by adjusting the unit's weights and/or threshold. In this way, they develop an effective mental model of neural computation that can be extended to more complex networks in later grades.

Students as AI creators

Our final curriculum design principle emphasizes the importance of students creating their own AI artifacts. We facilitate this via activities that introduce Scratch programming with AI extensions (Williams 2020), simulated intelligent robot programming with Calypso (Touretzky 2017), and the construction of classifiers using Teachable Machine, Machine Learning for Kids, and AI Lab.

TEACHER IMPACT

Through the co-design process, teachers reported becoming more knowledgeable about AI. In the words of one teacher who developed a better understanding about how AI works and its applications: "*I would say my cautious level (about AI) remains about the same… my understanding of how it works and what it's doing is far more informed*." They also developed a sense of ownership over the curriculum, feeling empowered to shape it in ways that met their students' needs and classroom cultures. For example, one teacher expressed her appreciation for the level of freedom afforded by co-design and how researchers trusted her expertise to make meaningful changes. This stood in sharp contrast with her previous teaching experiences:

"*…having been given information and told to teach this textbook just in the way that was given to me versus being invited to co-design… gives it a level of engagement… (and) opportunity to have a voice and a choice in it*."

Overall, teachers report that they benefit from the combination of professional learning before classroom implementation and support in the days leading up to classroom implementation, followed by reflection immediately afterward. Based on students' class work and responses, they changed their instructional approach, revised concepts, and suggested alternative problem-solving strategies. Teachers also modified their instruction based on post-implementation reflections of other teachers. Teachers helped make the material culturally relevant by connecting lesson content to real-world problems, to students' past experiences, to their interests, and to current events.

Two teachers from the first cohort have left the classroom and become educational consultants. They are still contributing to the project, helping to create more robust curriculum and teacher PD materials, and bite-sized offerings of the AI4GA curriculum modules.

STUDENT IMPACT

Prior to the first year of implementation, the research/evaluation team developed a pre- and post-survey to understand students' knowledge and attitudes towards artificial intelligence and previous experience with coding and computer programming. (The survey is included as supplementary material for this article.) Data from nearly 200 middle school students in the first 2 years of the project were analyzed to identify the impacts of the AI4GA curriculum. Our findings reveal that the project has had an impact on students' future orientation. The research also offers a contribution by revealing how middle school students' think about future careers and the impact of those thoughts on their attitudes toward AI.

After completing the AI4GA curriculum, students can list more AI careers, have an increased belief that their future career will involve artificial intelligence, and are more likely to believe that learning AI will help them get a good job someday. But despite teachers' efforts to deliver the content with fidelity, we found that students' interest in learning about AI and their comfort with learning about AI often declined from before they took the course to after. However, students who reported thinking about their future careers did not show this downward trend. Upon further exploration of the data, we hypothesize the following two factors that may contribute to this result: (1) Future orientation may allow students to get more out of the course, and (2) learning about the nuts and bolts of AI may conflict with students' prior conceptions and fundamentally change how they think about AI.

In this section, we discuss survey constructs and survey items, and it is important to understand what we mean. Each construct represents an idea or concept that cannot easily be measured directly, such as "Interest in Learning AI." Instead, we ask several related questions or items. For example, the construct "Interest in Learning AI" comprises items such as "I am curious about AI technologies" and "I think learning AI is relevant to my life."

A common goal for computing-related classroom interventions is shifting students' intentions to persist in STEM broadly and CS in particular, when thinking about their future adult selves. Conversations with other ITEST and AI education researchers reveal differing opinions on whether middle school students consider careers at this early age. To answer this question, we asked students to respond to the following two Likert scale statements on the pre/post survey: (1) I think about careers I might be interested in; (2) I think about problems I want to solve in the future. Students were given five response options as follows: strongly disagree, disagree, neutral, agree, and strongly agree. We collapsed student responses into three categories as follows: disagree (I do NOT think about careers), neutral, and agree (I DO think about careers). Table 4 reveals that of the 377 students included in our sample, over half (53.0%) of middle schoolers think about careers that they might be interested in. About one third (32.0%) of students were neutral, and 14.8% did not think about future careers.

The distribution in Table 4 is similar across demographic groups, both by gender and by race/ethnicity. That is, the demographic distribution of students in each group (DO NOT think about careers, neutral, and DO think about careers) is statistically similar (race/ethnicity: $\chi^2(10, N = 380) = 15.151$, $p = 0.127$; gender: $\chi^2(4, N = 380) = 6$, $p = 0.1991$).

Since the data for both questions are similar and space is limited, we focus our discussion on students' responses to the career item. Our findings reveal that whether or not students' think about their future careers influences their interest in learning AI, comfort with learning AI, attitudes towards AI technology, and thoughts about career decisions involving AI.

We found that students who do not consider careers experience significant pre-to-post declines across all construct areas, and those who are neutral experience significant declines across most construct areas (Table 5). Those who do think about careers start higher and either remain at that level or experience significant increases. Specifically, we found that students who do not think or are neutral about considering careers were significantly less interested in learning about AI at the end of the school

TABLE 4 Student perceptions of careers and problems they want to solve in the future.

Items	No	Neutral	Yes
I think about careers I might be interested in.	56 14.8%	121 32.0%	200 53.0%
I think about problems I want to solve in the future.	45 11.9%	143 37.9%	189 50.1%

TABLE 5 Change in student attitudes broken down by career orientation.

Construct	Career orientation	*n*	Mean Pre	Mean Post	*p*	Cohen's *D*
Interest in learning AI	I do NOT think about careers	56	3.30	2.45	<0.001***	0.887
	Neutral	121	3.66	3.36	<0.001***	0.0454
	I DO think about careers	200	3.89	3.81	0.173	0.101
Comfort with learning AI	I do NOT think about careers	56	3.44	2.77	<0.001***	0.601
	Neutral	121	3.90	3.56	<0.001***	0.462
	I DO think about careers	200	4.09	4.03	0.332	0.078
Attitudes toward AI technology	I do NOT think about careers	56	3.40	3.04	0.028*	0.354
	Neutral	121	3.61	3.60	0.871	0.017
	I DO think about careers	200	3.66	3.86	<0.001***	0.295
Thoughts about career decisions involving AI	I do NOT think about careers	56	2.91	2.20	<0.001***	0.816
	Neutral	121	3.29	3.17	0.033*	0.188
	I DO think about careers	200	3.47	3.69	<0.001***	0.256

Note: Response options: 1 = *strongly disagree*; 2 = *disagree*; 3 = *neither agree or disagree*; 4 = *agree*; 5 = *strongly agree*. * $p < 0.05$; *** $p < 0.001$. Asterisks indicate levels of statistical significance.

year. There was no change in interest in learning about AI for students who do think about careers. Students who do not think or are neutral about considering careers were significantly less comfortable with learning about AI at the end of the school year. There was no change in comfort with learning about AI for students who do think about careers. Students who do not think about careers had significantly less positive attitudes about AI technology at the end of the school year. Students who do think about careers had significantly more positive attitudes about AI technology at the end of the school year. Students who do not think or are neutral about their future careers had significantly less positive thoughts about career decisions involving AI at the end of the school year. Students who do think about careers had significantly more positive thoughts about career decisions involving AI at the end of the school year.

We disaggregated the data related to the item "My career someday might involve AI" (part of the "thoughts about career decisions" construct) by gender, and also by race/ethnicity using just two categories: groups that were historically underrepresented in STEM and groups that were historically well-represented. We found for all four groups (male, female, historically underrepresented in STEM, and historically well-represented in STEM), the AI4GA curriculum had a positive impact on their belief that their career someday might involve artificial intelligen ce (Table 6).

While the above analysis suggests a connection between considering future careers and increasing positive attitudes toward AI after taking the course, what could explain the decline in interest and comfort with learning about AI among students who don't think about future careers? We surmise that AI instruction may change students' thinking about AI in ways that instruction in other subjects (Mathematics, English/Language Arts, Social Science, etc.) may not. Students' prior exposure to AI in science fiction and popular culture, and their experiences with AI applications such as Alexa or ChatGPT, influence their preconceptions about and attitudes toward AI. The AI4GA curriculum exposes the technical details that undergird artificial intelligence systems (decision trees, word embeddings, neural networks, etc.), and these may not live up to the grand expectations students have when they enter the course. To account for students' shift in understanding, we changed our survey method and began collecting retrospective pre data during the third year of implementation. This allows students to reflect on their perceptions of AI before the program with the same understanding that they have after learning about AI.

TABLE 6 Anticipation of a future AI-related career by gender and ethnicity.

Construct	Groups	*n*	Mean Pre	Mean Post	*p*	Cohen's *D*
My career someday might involve AI.	Female	103	2.75	2.96	0.024*	0.176
	Male	92	3.03	3.30	0.006**	0.234
My career someday might involve AI.	Historically underrepresented	135	2.91	3.09	0.032*	0.148
	Historically well-represented	61	2.85	3.18	0.006**	0.269

Note: * p < 0.05; ** p < 0.01. Asterisks indicate levels of statistical significance.

TABLE 7 Comparing true pre/post to retrospective pre/post analyses by construct.

Construct	*n*	Pre condition	Mean Pre	Mean Post	*p*	Cohen's *D*
Interest in learning AI	198	True pre	3.67	3.37	0.002***	0.33
		Retrospective pre	3.36		0.939	0.01
Comfort with learning AI	197	True pre	3.83	3.59	0.02*	0.60
		Retrospective pre	3.62		0.691	0.46
Attitudes toward AI technology	195	True pre	3.56	3.52	0.95	0.04
		Retrospective pre	3.46		0.95	0.07
Thoughts about career decisions in AI	199	True pre	3.26	3.22	0.721	0.04
		Retrospective pre	3.11		0.51	0.11

Note: Response options: 1 = *strongly disagree*; 2 = *disagree*; 3 = *neither agree or disagree*; 4 = *agree*; 5 = *strongly agree*. * p < 0.05; *** p < 0.001. Asterisks indicate levels of statistical significance.

TABLE 8 Individual items showing notable differences by true and retrospective pre conditions.

Items	*n*	Pre condition	Mean Pre	Mean Post	*p*	Cohen's *D*
I can list careers that use AI.	199	True pre	3.13	3.29	0.067	0.148
	198	Retrospective pre	3.15	3.29	0.042*	0.117
I think learning about AI is relevant for my life.	196	True pre	3.33	3.27	0.518	0.052
	194	Retrospective pre	3.09	3.26	0.012*	0.150
My career someday might involve AI.	197	True pre	3.14	3.13	0.909	0.009
	199	Retrospective pre	2.89	3.13	<0.001***	0.197
Learning about AI will help me get a good job someday.	195	True pre	3.42	3.34	0.314	0.074
	196	Retrospective pre	3.19	3.34	0.012*	0.125

Note: Post values may be different for true pre and retrospective pre conditions to reflect the different number of matched responses. For example, for the item "I think learning about AI is relevant for my life," two students did not respond to the retrospective pre items. Response options: 1 = *strongly disagree*; 2 = *disagree*; 3 = *neither agree or disagree*; 4 = *agree*; 5 = *strongly agree*. * p < 0.05; *** p < 0.001. Asterisks indicate levels of statistical significance.

We conducted a retrospective pre-survey where students were asked at the end of the course to imagine how they felt prior to the course: students were reflecting back on their prior state (retrospective). Table 7 shows differences in pre condition (true pre vs. retrospective pre) by construct and shows that interest and comfort significantly decline in the true pre to post condition but remain statistically similar in the retrospective pre to post condition. In other words, at the end of the course students underestimate their prior levels of interest and comfort in learning about AI. On the other hand, Table 8 shows notable items in which the retrospective condition shows statistically significant increases while the true pre condition does not. So for these items, students report that their agreement with the item increased after taking the course, when in reality it did not. This is not uncommon as students become more familiar with a topic of interest. DiSalvo and Bruckman (2009) and DiSalvo et al. (2014) found that teaching Black male high school students who liked playing games how to test games, decreased some students' interest in games and did not increase their interest in taking subsequent CS courses or pursuing CS careers.

While the student survey results are mixed, the retrospective results show that students' high interest and comfort with learning AI remains high (see Table 7). Further, students' attitudes toward AI technology and thoughts on AI career decisions significantly increase for those students indicating they think about careers (see Table 5). Additionally, we note a significant shift in the way students respond to items on the pre condition of the survey compared to the retrospective (reflecting back) condition, indicating that AI instruction may fundamentally change the way students think about AI. This shift may diminish as AI is infused into schooling. The next step in our research is to better understand how initial perceptions and definitions of AI are changing as students become more accustomed to learning alongside AI agents.

SUSTAINABILITY AND EVOLUTION

From its inception, a major goal in the design of the AI4GA project has been sustainability. We designed for redundancy in role coverage across the team, as well as diversity of skill sets. We have designed the curriculum and PD program to be easily adapted based on the expertise and instructional styles of the teachers. Our goal is to ensure that we are building infrastructure that will outlive any shifts in project team members or funding. At the core of the infrastructure plan are ongoing PD opportunities, development of middle school AI teacher leaders, an AI teacher professional learning community, and state course standards for an AI middle school elective. At the time of this writing, we are on our third stream of funding for this project, and we have had several personnel shifts at key positions, but the project still thrives.

We are currently leveraging our infrastructure model in a new project funded by Google: AI for Middle Schools; —Multi-State Scale-up. This project is allowing us to partner with other CS Education organizations in Texas and Florida to provide introductory teacher PD and teacher leader training. Our goal is to train 25 new teachers in Georgia, Florida, and Texas in Summer 2025. Leveraging our partnership with WeTeach_AI Champions in Austin, TX, we are excited to be training our second cohort of teacher leaders who will enable Texas to continue to offer PD beyond their partnership with us.

GaDOE online AI teacher short courses

By the end of the second year, the first cohort of teacher leaders pulled together their experiences being taught about and then teaching AI to create a series of asynchronous online short courses, each running from 1 to 2 h. The courses are hosted on the Georgia Department of Education's professional learning platform and are available to all K-12 teachers in the state. The courses were not intended to be a comprehensive professional learning experience but rather a first step to understanding what AI education in K-12 looks like.

There are currently three courses in the series, each corresponding to a unit in the AI4GA curriculum: Autonomous Robots and Vehicles, How Computers Understand Language, and How Computers Make Decisions. The courses can be used by teachers aspiring to teach AI in middle school, teachers who want a foundational understanding of AI concepts to integrate them into other content areas, administrators who are interested in adding AI to their school offerings, counselors, media specialists, and anyone else who would benefit from a fundamental AI literacy. Each course includes interactive activities and external links for demos, additional support, and further reading.

As of April 2025, 79 educators have taken the first course (Autonomous Robots and Vehicles), 126 have taken the second (How Computers Understand Language), and 137 people have taken the third (How Computers Make Decisions).

AI explorer guild

The AI4GA project was instantiated in the midst of the COVID epidemic with year 1 being 2021–2022. The project was well underway when ChatGPT took the world by storm in November of 2022. With the advent of LLMs, interest in the use of AI in K-12 exploded. Although the AI4GA project was exploring AI concepts and not necessarily the use of AI tools in education, the project team, including the middle school teachers, still found themselves in a fortuitous position, able to speak to the underpinnings of AI in a way that was accessible to other educators. This dialogue, including discussion about the differentiation between teaching with AI and teaching about AI, began to happen in school hallways, Board of Education meetings, core content area conferences, and every other space where educators were able to reflect on the impact AI was having on education.

The AI4GA team saw a need to create a space where educators could have these conversations, ask the pressing questions, with an educator in the room who could speak to some of their queries and concerns. The Georgia Tech-based PI, now working at the Constellations Center for Equity in Computing, collaborated with the K-12 outreach team at Georgia Tech Research Institute to create a regular virtual meetup, providing a space for educators to convene and discuss the advent of AI in the K-12 space.

These meetings would start with a brief presentation from a researcher or an AI4GA teacher, and then shift into an open discussion and Q&A session. The only guideline is that the discussions have to be about teaching with or teaching about AI. The series began in the Fall of 2024 and continues into the Spring of 2025.

Expanding middle school AI across disciplines

At the start of the project, we focused exclusively on recruiting CS teachers who were already teaching other middle school computing courses. We are now broadening our outreach to include teachers in other disciplines who are interested in incorporating some aspects of AI into their curriculum. One example is a teacher leader from our third cohort who teaches a sixth-grade science class. She adapted our Candy Land decision tree activity to work with her rocks and minerals unit. Her students create mineral classification decision trees, leveraging the similarities between decision tree classifiers in AI and “dichotomous keys” used in the natural sciences (National Park Service n.d.). This teacher is now helping to train other teachers at her school to adapt bits of AI curriculum for their own disciplines.

Our first teacher PD event in Texas, in July 2024, included teachers from many different disciplines, including Science, Mathematics, English and Language Arts, CS, Robotics, Applied Engineering, Technology, App Design, Professional Communication, and Career Exploration. We are continuing to explore ways to support teachers in incorporating AI concepts into their disciplines. For this work, we plan to leverage the best practices and models used to infuse computational thinking into Science, Math, English, and Social Studies (such as Jocius et al. 2020)

Contributions to high school AI education

The AI4GA team has been engaged in several spin-off projects based on the AI4GA curriculum. The BridgeUP STEM program, led by Judith Uchidiuno, engages high school students in an after-school program for the creation and evaluation of educational games that teach the five big ideas available at AI4K12.org. The project uses the AI4GA curriculum to provide a foundation for their knowledge during this program (Lim et al. 2025). This program highlights the benefits of co-design for designing AI education games, incorporating role playing and competitive game mechanics to improve student engagement, uncovers students’ AI misconceptions in authentic settings, and showcases students’ ability to apply AI concepts in game-based learning.

One of the rural districts we have partnered with is also looking to link the middle school AI course to their high school AI CTAE pathway. We are looking to replicate this model in other districts that are teaching either our middle school AI course or the approved high school AI pathway. Based on our evaluation data, we anticipate that the middle school students who are interested in AI careers will want a next step for learning AI in high school.

Plans for future curriculum development

The AI4GA curriculum continues to evolve. We plan to expand our coverage of large language models, looking not only at prompt engineering but also at their reasoning abilities and their prospects for achieving artificial general intelligence (AGI). Experimentation with LLMs in middle school is complicated by the fact that some school districts block access to these systems, and many providers restrict accounts to users age 13 or older. Given the growing contribution of LLMs to everyday life and to many lines of work, we feel it is important to find ways to overcome these obstacles.

Another area of planned innovation is the introduction of AI-powered robots. Students have on numerous occasions expressed a desire to not just read about autonomous robots, but to experience them hands-on. The demise of the Cozmo robot with the 2019 bankruptcy of its manufacturer, Anki, left no good options until recently, when Innovation First announced the VEX AIM. We are currently conducting preliminary experiments on interfacing VEX AIM with GPT-4o.

Our original outline of the AI4GA curriculum had five units instead of three. We have a partial draft of Unit 4, which is devoted to intelligent agents. Unit 5 was to cover careers in AI and robotics. Although five units proved to be too much to fit into a 9-week course, we will continue to develop these last two units to provide teachers with more choices to select from. One early resource from Unit 5, a collection of AI career cards, is already available on the AI4GA.org website.

ACKNOWLEDGMENTS

We thank our colleagues, Lavita Williams and Alba Guttierez, at the Georgia Department of Education for their many contributions to this project. “AI for Georgia” was funded by National Science Foundation awards DRL-2049029 and DRL-2048502, and by gifts from Google.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict.

ORCID

David S. Touretzky https://orcid.org/0000-0002-9388-4970

REFERENCES

Almeda, M. V., and R. S. Baker. 2020. "Predicting Student Participation in STEM Careers: The Role of Affect and Engagement during Middle School." *Journal of Educational Data Mining* 12(2): 33–47. https://doi.org/10.5281/zenodo.4008054.

Bandyopadhyay, S., J. Xu, N. Pawar, and D. Touretzky. 2022. "Interactive Visualizations of Word Embeddings for K-12 Students." *Proceedings of the AAAI Conference on Artificial Intelligence* 36(11): 12713–20. https://doi.org/10.1609/aaai.v36i11.21548.

Computer Science Teachers Association. 2017. "CSTA K–12 Computer Science Standards." Revised 2017. https://csteachers.org/k12standards.

CSforALL. 2024. "The SCRIPT: Strategic CSforALL Planning Tool for School Districts." https://www.csforall.org/projects_and_programs/script (accessed March 19, 2025).

DiSalvo, B. J., and A. Bruckman. 2009. "Questioning Video Games' Influence on CS Interest." In *Proceedings of the 4th International Conference on Foundations of Digital Games (FDG'09)*, 272–78. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/1536513.1536561

DiSalvo, B., M. Guzdial, A. Bruckman, and T. McKlin. 2014. "Saving Face While Geeking Out: Video Game Testing as a Justification for Learning Computer Science." *Journal of the Learning Sciences* 23(3): 272–315. https://doi.org/10.1080/10508406.2014.893434.

Google Research. 2018. "Semantris." https://research.google.com/semantris (accessed March 9, 2025).

Gelder, W., X. Yu, D. Touretzky, C. Gardner-McCune, and J. Uchidiuno. 2025. "From Lecture Hall to Homeroom: Co-Designing an AI Elective with Middle School CS Teachers." *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-024-00449-3.

Goode, J., J. Margolis, and G. Chapman. 2014. "Curriculum is Not Enough: The Educational Theory and Research Foundation of the Exploring Computer Science Professional Development Model." In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE'14)*, 493–98. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2538862.2538948.

Henrick, E., P. Cobb, W. Penuel, K. Jackson, and T. Clark. 2017. "Assessing Research-Practice Partnerships: Five Dimensions of Effectiveness." William T. Grant Foundation. https://wtgrantfoundation.org/library/uploads/2017/10/AssessingResearch-Practice-Partnerships.pdf.

Jocius, R., D. Joshi, Y. Dong, R. Robinson, V. Cateté, T. Barnes, J. Albert, A. Andrews, and N. Lytle. 2020. "Code, Connect, Create: The 3C Professional Development Model to Support Computational Thinking Infusion." In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE'20)*, 971–77. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3328778.3366797.

Lee, I., S. Ali, H. Zhang, D. DiPaola, and C. Breazeal. 2021. "Developing Middle School Students' AI Literacy." In *Proceedings of SIGCSE'21*, 191–97. https://doi.org/10.1145/3408877.3432513.

Lim, J., G. Barkhuff, J. Awuah, S. Clyde, R. Sogani, D. Touretzky, C. Gardner-McCune, and J. Uchidiuno. 2025. "Escape or D13: Understanding Youth Perspectives of AI through Educational Game Co-design." In *Proceedings of the 40th Annual ACM Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3714037.

National Park Service. n.d. "Using Dichotomous Keys." https://www.nps.gov/teachers/classrooms/dichotomous-key.htm (accessed March 9, 2025).

McGill, M. M., A. Peterfreund, S. Sexton, R. Zarch, and M. Kargarmoakhar. 2021. "Exploring Research Practice Partnerships for Use in K–12 Computer Science Education." *ACM Inroads* 12(3): 24–31. https://doi.org/10.1145/3477607.

Phillips, K. 2019. "Teachable Machine 2.0 Makes AI Easier for Everyone." Google Blog. https://blog.google/technology/ai/teachable-machine (accessed April 8, 2025).

Price, T. W., V. Cateté, J. Albert, T. Barnes, and D. Garcia. 2016. "Lessons Learned from "BJC" CS Principles Professional Development." In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE'16)*, 467–72. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2839509.2844625.

Rhode Island Department of Education & National Center for the Improvement of Educational Assessment. n.d. "Student Work Analysis Protocol." https://ride.ri.gov/sites/g/files/xkgbur806/files/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/Educator-Evaluation/Online-Modules/Student_Work_Analysis_Protocol.pdf (accessed April 8, 2025).

Rosato, J., C. Lucarelli, C. Beckworth, and R. Morelli. 2017. "A Comparison of Online and Hybrid Professional Development for CS Principles Teachers." In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE'17)*, New York, NY: Association for Computing Machinery, 140–45. https://doi.org/10.1145/3059009.3059060.

Roschelle, J., and W. R. Penuel. 2006. "Co-design of Innovations with Teachers: Definition and Dynamics." In *Proceedings of the 7th International Conference on Learning Sciences (ICLS'06)*, 606–12. International Society of the Learning Sciences.

Touretzky, D. S. 2017. "Computational Thinking and Mental Models: From Kodu to Calypso." In *2017 IEEE Blocks and Beyond Workshop (B&B)*, 71–78. Raleigh, NC. https://doi.org/10.1109/BLOCKS.2017.8120416

Touretzky, D., C. Gardner-McCune, C. Breazeal, F. Martin, and D. Seehorn. 2019. "A Year in K-12 AI Education." *AI Magazine* 40(4): 88–90, Winter 2019. https://doi.org/10.1609/aimag.v40i4.5289.

Touretzky, D., C. Gardner-McCune, F. Martin, and D. Seehorn. 2019. "Envisioning AI for K-12: What Should Every Child Know About AI?" In *Proceedings of AAAI-19*, Honolulu, HI, January, 2019. https://doi.org/10.1609/aaai.v33i01.33019795.

Touretzky, D. S., A. Chen, and N. Pawar. 2024. "Neural Networks in Middle School." *ACM Inroads* 15(3): 24–28. https://doi.org/10.1145/3677610.

Touretzky, D., C. Gardner-McCune, W. Hanna, A. Chen, and N. Pawar. 2025. "Learning to Think Like a Neuron in Middle School."

In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39. https://doi.org/10.1609/aaai.v39i28.35195.

Trotman, A. 2017. "Why Don't European Girls Like Science or Technology?" Blog post of March 1, 2017 https://news.microsoft.com/europe/features/dont-european-girls-like-science-technology (Accessed May 20, 2025).

Trump, D. J. 2025. "Advancing Artificial Intelligence Education for American Youth (Executive Order No. 14179)." The White House. 90 FR 17519. April 23, 2025. https://www.federalregister.gov/d/2025-07368

Williams, R. 2020. "RAISE AI Playground." Blog post at https://www.media.mit.edu/posts/ai-blocks (accessed April 6, 2025).

Yu, X., W. Gelder, D. Touretzky C. Gardner-McCune, and J. Uchidiuno. 2025. "Design Considerations for Evaluating Middle School AI Knowledge." In *International Conference of the Learning Sciences 2025*, 1978–82. Helsinki, Finland, June 10–13, 2025.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Touretzky, D. S., C. Gardner-McCune, B. Cox, J. Uchidiuno, X. Yu, W. Gelder, T. McKlin, T. L. Brown, B. Kareem, W. Chung, A. Jones, and J. Kolodner. 2025. "*Artificial Intelligence Education in Georgia Middle Schools*." *AI Magazine* 46: e70013.https://doi.org/10.1002/aaai.70013

AUTHOR BIOGRAPHIES

David S. Touretzky is a research professor of Computer Science at Carnegie Mellon University and founder and chair of AI4K12.org.

Christina Gardner-McCune is an associate professor in the Computer & Information Science & Engineering Department at the Herbert Wertheim College of Engineering at the University of Florida. She is also a co-founder and the co-chair of AI4K12.org.

Bryan Cox is a research faculty at the Constellations Center for Equity in Computing/College of Computing at the Georgia Institute of Technology and is also a research fellow for the Kapor Center and is responsible for organizing the CSforAtlanta initiative.

Judith Uchidiuno is an assistant professor at Georgia Institute of Technology's School of Interactive Computing and directs the Play and Learn Lab.

Xueru Yu is a Ph.D. student in Human-Centered Computing at Georgia Institute of Technology's School of Interactive Computing.

William Gelder is a research associate of the Play and Learn Lab at the Georgia Institute of Technology.

Tom McKlin is the director of the The Findings Group.

Taneisha Lee Brown is a lead evaluator for The Findings Group.

Bejanae Kareem is the executive director at BK International Education Consultancy, specializing in multigenerational STEM programs, program coordination, and professional learning initiatives.

Woojin Chung is a data analyst specializing in business intelligence and data-driven storytelling.

Amber Jones is a former middle school computing teacher who now offers educational consulting services through *Amber Sparks Education*.

Janet Kolodner is an emeritus professor at Georgia Institute of Technology's School of Interactive Computing and Boston College.