Catching the First Light of Tomorrow: A Hackathon-Based Framework for Introducing High School Students to AI Agents

Long Nguyen[1], Duc Nguyen[1], Quan Bui[1], Dung Phan[1], Dung Le[1], Khanh Nguyen[1], Quynh Vo[1], Khang Vo[2], Nam Duong[1], Anh Dinh[1], Tri Trinh[1], Chi Phan[1], An Nguyen[1], Thai Nguyen[1], Dang Le[1], Vinh Dang[3], Tho Quan[1*]

[1]URA Research Group, Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Vietnam
[2]Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates
[3]British University Vietnam, Vietnam
{long.nguyencse2023, qttho}@hcmut.edu.vn

Abstract

Artificial Intelligence (AI), particularly in the form of intelligent AI agents, is transforming education, industry, and everyday life. These agents extend the capabilities of Large Language Models (LLMs) by integrating planning, decision-making, tool use, and multi-agent collaboration, enabling systems that can reason, adapt, and act in dynamic environments. As such systems become integral to modern workplaces and everyday problem solving, early exposure equips high school students with systems thinking, practical problem-solving skills, and ethical awareness, while preparing them to create applications that address real-world needs. Yet most high school AI programs focus on basic model usage and overlook the skills required to design and deploy agentic systems. Existing resources are largely aimed at university learners and assume substantial programming expertise, creating a significant accessibility gap. To address this need, we present a structured hackathon-based framework for introducing high school students to the design and application of AI agents. The framework combines expert-led lectures on core topics such as agent architectures, prompting strategies, reasoning methods, and tool-use protocols with a guided hackathon in which students collaboratively develop domain-specific agent-based chatbots. We provide a complete suite of instructional materials, step-by-step tutorials, and starter code to support hands-on learning, enabling participants to build functional agents capable of reasoning and interacting with external tools. Our approach bridges the gap between AI literacy and practical deployment while fostering creativity, collaboration, and responsible innovation, and our findings suggest that early engagement with agent-based AI design equips students with both technical proficiency and the mindset to shape the AI-driven future.

Supplementary Materials — https://osf.io/vx7gu/

1 Introduction

The rapid advancement of *Artificial Intelligence* (AI) has reshaped nearly every domain, from education and communication to scientific discovery and automation (Tan, Cheng, and Ling 2025). As intelligent technologies become increasingly embedded in daily life, equipping students with the knowledge and skills to understand and engage with AI systems has emerged as a critical goal of modern education. The concept of *AI literacy*, which refers to the ability to understand, apply, and critically evaluate AI systems, is now widely recognized as a core competency for the twenty-first century (Long and Magerko 2020; Nguyen et al. 2025a). In response, school systems around the world have begun integrating AI instruction into K–12 curricula, with particular momentum at the high school level (Alonso 2020; Norouzi, Chaturvedi, and Rutledge 2020; Grover 2024). At this formative stage, students are not only building academic foundations but also exploring long-term interests and potential career paths. Preparing them for an AI-driven future requires moving beyond superficial tool usage and fostering the capacity to design, analyze, and reason about intelligent systems (Touretzky et al. 2019).

The emergence of *Large Language Models* (LLMs) has significantly expanded access to AI through intuitive natural language interfaces. From chatbots and content generators to virtual assistants, LLMs now power a wide range of applications and serve as accessible entry points to advanced technologies (Sajja et al. 2024). While this accessibility has sparked broad interest, it also introduces new educational challenges. Teaching students to prompt pre-trained models may offer an initial foothold, but it does not provide the depth needed to understand how intelligent systems function (Varuvel Dennison et al. 2024). The next frontier in AI education lies in agentic systems, which combine LLMs with capabilities such as planning, memory management, decision-making, and tool use. These agents can retrieve external knowledge, sustain multi-turn dialogues, invoke APIs, and perform goal-directed tasks. Rather than functioning as isolated models, they act as orchestrators that coordinate reasoning and execution across modular components and dynamic environments (Maldonado et al. 2024). Such agentic capabilities now underpin a growing class of real-world systems, including personal copilots, domain-specific assistants, and retrieval-augmented solvers in education, healthcare, and beyond.

Empowering students to participate meaningfully in this paradigm shift requires more than technical exposure. It calls for cultivating a mindset that reflects the core characteristics of intelligent agents, including modularity, reasoning, adaptability, and collaboration. Students must learn how AI systems are structured, how they make decisions, and how multiple agents can work together to solve complex tasks. By engaging with concepts from multi-agent systems, learners begin to appreciate how intelligent behavior can emerge from distributed reasoning and interaction among autonomous components. This perspective also opens the door to discussions of transparency, responsibility, and ethical alignment, which are critical themes as AI increasingly shapes societal outcomes (Wiese et al. 2025; Russell and Norvig 2009). Despite these needs, the agent-based perspective remains largely absent from current high school curricula. Educational resources on AI agents are typically designed for university students and working professionals, often offered through undergraduate programs or online platforms such as Coursera[1] and Udemy[2]. These offerings generally assume significant prior knowledge in programming and engineering, making them less accessible to younger learners. Most existing high school programs focus on general AI awareness or introductory machine learning, with limited attention to modular system design, decision-making under uncertainty, or goal-oriented behavior (Lim et al. 2024; Olari, Cvejoski, and Eide 2021; Broll and Grover 2024). In addition, many materials are not tailored to the unique pedagogical needs of adolescent students.

To address this gap, educators need approaches that are both pedagogically engaging and technically rigorous (Wang et al. 2024; Yim and Su 2025; Rizvi, Waite, and Sentance 2023). Active learning strategies, such as collaborative projects, inquiry-based exploration, and hands-on experimentation, have proven effective in STEM education. Among these, hackathons offer a particularly promising format. By immersing students in time-bounded, team-oriented challenges, hackathons promote deep engagement with technical content while also fostering essential skills such as creativity, communication, and critical thinking (Surendran et al. 2023; Araujo, Kalinowski, and Baldassarre 2024; Nguyen et al. 2025a). Several prominent initiatives, such as the Regeneron *International Science and Engineering Fair*[3] (ISEF) and the *International Olympiad in Artificial Intelligence*[4] (IOAI), have introduced high school students to AI. However, these efforts often emphasize algorithmic problem solving or pre-trained model usage, with limited attention to the design and deployment of complete agentic systems. This gap highlights the need for educational experiences that guide students in building AI systems from an agent-centric and ethically informed perspective.

In this paper, we present a structured hackathon-based framework designed to introduce high school students to the foundations and applications of AI agents. Our framework consists of two core components. First, students participate in a series of expert-led, interactive lectures covering foundational topics such as generative AI, agent architectures, prompting strategies, reasoning techniques, and tool-use protocols. These lectures are carefully scaffolded to make complex ideas accessible while preserving technical depth. Second, students engage in a guided hackathon in which they collaboratively design and implement agent-based chatbots tailored to real-world domains such as education and healthcare. These systems are expected to reason over tasks, interact with tools, and respond dynamically to user goals. Our contributions are summarized as follows.

- We develop a comprehensive suite of educational materials, including lecture slides and instructional videos, for teaching the foundations of AI agents and their integration into real-world systems.
- We provide detailed tutorials with step-by-step guidance and starter code, enabling students to build functional agent-based chatbots capable of reasoning and interacting with tools.
- We propose a structured hackathon framework specifically designed for high school learners, bridging the gap between AI literacy and practical deployment through guided, domain-relevant chatbot design.

2 Related Works

2.1 Contemporary Approaches to Teaching AI in High School Education

High school AI education has evolved from abstract theory or basic tool familiarization toward scaffolded, project-based learning that links conceptual understanding with artifact creation (Sajja et al. 2024; Wang et al. 2024). At the framework level, the AI4K12 initiative outlines the *Five Big Ideas*[5] and grade-band progressions for Grades 9–12, while programs such as Day of AI[6] and Experience AI[7] provide lesson plans, classroom activities, and ethics modules. Empirical studies show that complex AI topics can be taught effectively with appropriate scaffolding. Examples include explainable-AI workshops using Scratch classifiers to highlight bias and interpretability (Alonso 2020), month-long *Machine Learning* (ML) and *Natural Language Processing* (NLP) programs introducing Python and curated datasets (Norouzi, Chaturvedi, and Rutledge 2020), and guided LLM environments cultivating structured metacognitive prompting (Varuvel Dennison et al. 2024). Complementary unplugged and low-ceiling/high-ceiling activities broaden access, from AI-themed board games (Lim et al. 2024) and block-based robotics curricula (Olari, Cvejoski, and Eide 2021) to interactive exercises on optimization and adversarial examples (Broll and Grover 2024). Hackathon-style models have also been piloted, combining structured tutorials with culminating challenges (Nguyen et al. 2025a), highlighting the value of guided, time-bounded learning. Across these initiatives, three principles recur: explicit scaffolds, authentic tasks with public artifacts, and barrier-lowering tools that democratize participation. However, most approaches stop short of agentic systems; few offer end-to-end support

[1]https://www.coursera.org/
[2]https://www.udemy.com/
[3]https://www.societyforscience.org/isef/
[4]https://ioai-official.org/
[5]https://ai4k12.org/
[6]https://dayofai.org/
[7]https://experience-ai.org/

for planning, memory, external tool use, or collaborative workflows. This gap motivates our hackathon-based framework, extending prior project-based efforts toward functional agentic systems for high school learners.

2.2 Educational Competitions Focused on AI Agents for High School Students

Competitions and hackathons offer authentic and motivating opportunities for students to apply AI knowledge in practice (Surendran et al. 2023; Araujo, Kalinowski, and Baldassarre 2024; Nandi and Mandernach 2016; Zhang, Wu, and Yu 2021). Existing initiatives such as IOAI and Microsoft Imagine Cup Junior[8] broaden participation but primarily emphasize general applications (e.g., image recognition, text analysis, ethical reflection) rather than the design of autonomous agentic systems. Consequently, students often gain conceptual awareness or experience with prebuilt tools without developing agents capable of planning, decision-making, and context-sensitive interaction. More recently, industry-led events such as the Microsoft AI Agents Hackathon[9], the LabLab–MindsDB AI Agents Hack[10], and the NVIDIA Agent Toolkit Hackathon[11] have emphasized agentic capabilities through starter templates, expert-led tutorials, and evaluation criteria highlighting planning, autonomous execution, and adaptive interaction. However, these events are largely oriented toward developers, university students, and professionals; when high school students are eligible, the absence of age-appropriate scaffolding and sustained instructional support creates substantial entry barriers. This underscores the need for dedicated frameworks that bring agentic AI—particularly agent-based chatbots—into high school education. Our framework addresses this gap by integrating expert-led mini-lectures on agent reasoning, live demonstrations, team-based mentorship, and a structured handbook with starter code, checklists, and troubleshooting guidance tailored to high school learners.

3 An AI Agent Hackathon-based Educational Framework

3.1 Learning Outcomes

Our hackathon-based framework adopts *Bloom's Taxonomy* (Bloom 1956) as the pedagogical backbone for defining learning outcomes and assessments. Originally introduced by (Bloom 1956) as a six-level classification of cognitive objectives, the taxonomy was later revised by (Anderson and Krathwohl 2001) into processes (*Remember* → *Understand* → *Apply* → *Analyze* → *Evaluate* → *Create*), which better align with contemporary educational practices. Guided by this framework, we designed six learning outcomes (LO1–LO6) that collectively span the full cognitive spectrum.

LO1 – Identify and describe fundamental concepts related to AI agents.

- **LO1.1** – Recall definitions of *AI agent*, *prompt*, and *multi-agent system*.
- **LO1.2** – Differentiate between traditional software and agentic AI.
- **LO1.3** – Identify key terminology related to agent architectures.

LO2 – Explain core processes of AI agents such as perception, reasoning, and learning.

- **LO2.1** – Explain the roles of perception, reasoning, and learning in AI systems.
- **LO2.2** – Provide real-world examples illustrating these processes.

LO3 – Apply AI tools to accomplish domain-specific tasks.

- **LO3.1** – Implement a basic AI agent pipeline using provided tools.
- **LO3.2** – Apply prompt engineering techniques to complete assigned tasks.

LO4 – Analyze and propose AI prompts and system workflows, identifying strengths, weaknesses, and trade-offs.

- **LO4.1** – Propose and compare alternative prompts or workflows for a given task.
- **LO4.2** – Break down agent roles, architectural choices, and interactions in multi-agent systems.
- **LO4.3** – Assess potential impact and feasibility of different workflow designs.

LO5 – Evaluate AI system outputs and justify design or tool choices with evidence.

- **LO5.1** – Critique the outputs of an AI agent in terms of accuracy and relevance.
- **LO5.2** – Justify model or tool choices in presentations using evidence.

LO6 – Design and develop innovative AI solutions or workflows to address real-world challenges.

- **LO6.1** – Formulate a project idea addressing an educational or social problem.
- **LO6.2** – Design the system architecture and workflow aligned with the proposed impact.
- **LO6.3** – Develop and present a functional prototype with a clear use case.

This framework ensures that students progress systematically from foundational knowledge acquisition (LO1–LO2), to practical application (LO3), higher-order reasoning (LO4–LO5), and ultimately innovation (LO6), in alignment with principles of outcome-based education (Spady 1995).

To maintain alignment across objectives, instruction, and assessment, the outcomes were operationalized through targeted learning opportunities embedded in the hackathon,

[8]https://imaginecup.microsoft.com/

[9]https://microsoft.github.io/AI_Agents_Hackathon/

[10]https://lablab.ai/event/ai-agents-hack-with-lablab-and-mindsdb

[11]https://developer.nvidia.com/agentiq-hackathon

LO	Performance Metrics	Example Learning Opportunities
LO1	- Accuracy on terminology quizzes - Clarity in defining concepts	- Expert-led introductory lectures on AI agents - Reading materials and glossary tasks
LO2	- Quality of explanations in short answers - Peer feedback on clarity	- Live demonstrations with expert Q&A - Guided group discussions
LO3	- Successful completion of assigned tasks - Functionality of implemented outputs	- Hands-on coding tutorials with provided tools - Mentor-supported challenges
LO4	- Ability to compare and refine alternative prompts/workflows - Depth of analysis of agent roles and interactions	- Group activities analyzing and redesigning workflows - Mentor-led design critique sessions
LO5	- Justification of design decisions - Quality of evaluation criteria applied	- Structured mentor-guided feedback loops - Peer review of outputs during hackathon
LO6	- Creativity, feasibility, and impact of final project - Integration of multiple system components	- Team-based hackathon sprints with mentor support - Final public presentation and demonstration

Table 1: *Learning Outcome* (LO) metrics and associated learning opportunities for the AI Agent Hackathon.

including expert-led lectures, guided exercises, collaborative project work, expert feedback sessions, and mentor-led sprints. Student progression was measured using parallel pre- and post-test instruments, each with six sections of five questions mapped to Bloom's cognitive levels. This structure captured both surface-level knowledge and higher-order competencies such as problem-solving, evaluative reasoning, and creative synthesis. As summarized in Table 1, each learning outcome corresponds to measurable performance indicators and authentic learning activities: expert-led lectures and demonstrations (LO1–2), mentor-supported exercises and mini-challenges (LO3), group analysis and expert feedback (LO4–5), and team-based hackathon projects culminating in public presentations (LO6). Together, these outcomes form the pedagogical backbone of our framework, guiding instructional design and hackathon development.

3.2 Instructional Materials

Overview The instructional materials were carefully designed as a multi-layered pedagogical scaffold supporting the hackathon. Rather than functioning as isolated resources, the package integrates expert-led lectures, multimedia supplements, a structured tutorial handbook, and an open-source codebase. Together, these components form a coherent learning pathway that balances conceptual grounding with incremental practice, enabling students to progress from introductory exposure to the design and deployment of functional multi-agent systems. To ensure transparency and reproducibility, the full set of materials is available in our archive at https://osf.io/vx7gu/.

Slides and Videos The lecture series was sequenced to build conceptual foundations while remaining accessible to high school learners. Each session combined clear explanations, live demonstrations, and interactive Q&A to foster active engagement.

- **Lecture 1: From AI to Agentic AI.** This session traced the evolution of AI from rule-based systems and ML to generative models and multi-agent paradigms. It emphasized three core capabilities—perception, reasoning, and learning—and showed how they extend through agentic workflows such as task decomposition, memory management, and adaptive planning. The lecture concluded with a case study of our university's enrollment consultation chatbot to ground key ideas in a real-world system.
- **Lecture 2: Agentic AI and Multi-Agent Systems.** This lecture introduced defining characteristics such as autonomy, adaptability, and goal-directed behavior. It provided an overview of architectural modules (perception, memory, planning, action), organizational structures (centralized, decentralized, hybrid), and communication protocols (e.g., MCP, agent-to-agent messaging). Applications in healthcare, climate action, and smart cities highlighted both societal relevance and technical depth.

Tutorial Handbook Complementing the lectures, the tutorial handbook offered a scaffolded progression from conceptual framing to hands-on implementation. Organized into sequential chapters, it guided students through environment setup, building a simple multi-agent system, constructing small datasets, and deploying modular architectures. The handbook also introduced the open-source codebase and illustrated how to extend it to domains such as education, healthcare, and smart cities. Each chapter combined visual walkthroughs, simplified scripts, and mentor guidance to ensure accessibility and reinforce both conceptual understanding and technical proficiency.

Demonstration A live demonstration of the agentic enrollment consultation chatbot provided a practical bridge between theory and implementation. By illustrating workflow design, feasibility analysis, and system-level evaluation, the demonstration offered both inspiration and a concrete benchmark for student projects.

Codebase The open-source codebase served as the technical backbone of the hackathon. Its modular design included ready-to-use tools, agent templates, and orchestration utilities, documented to minimize barriers for beginners. To support rapid onboarding and deeper exploration, the package provided a quick-start guide, UI customization notes, and integration examples for creating new agents, using tools, and handling complex queries. This structure enabled participants to focus on creativity, workflow design, and impact assessment rather than technical setup.

3.3 Pre-Implementation Preparation

Mentor Teams. Assigning qualified mentors to each student team was a central element of our design. We recruited senior undergraduates and graduate students using a weighted composite score, defined in Equation (1), which balanced academic foundation, research involvement, and industry exposure:

$$S = 0.5F + 0.3R + 0.2I. \tag{1}$$

Here, F denotes the average grade in core courses such as *AI*, *Intelligent Systems*, and *NLP*, normalized to a 10-point scale ($A^+ = 10$, $A = 9$, $A^- = 8$, $B^+ = 7.5$, $B = 7$, $B^- = 6.5$). R and I represent months of active research engagement and industry experience, respectively, each rescaled to a maximum of 10. Selected mentors completed a two-week training program combining conceptual reinforcement with hands-on agentic AI practice, preparing them to guide participants effectively throughout the hackathon.

Invited Speakers. To broaden perspectives and inspire participants, we invited three distinguished speakers whose contributions complemented the instructional program. Two professors with expertise in AI and multi-agent systems delivered lectures that established conceptual foundations and introduced deployment strategies. In parallel, the Chief Technology Officer of a leading technology company presented a live demonstration of an admission-consulting chatbot, illustrating real-world relevance. Together, these sessions connected theory with practice and highlighted the societal impact of agentic systems.

Awards and Prizes. A structured award system was introduced to recognize excellence and sustain motivation. Prizes were organized into First, Second, and Third Prize categories, along with two Honorable Mentions. Depending on available resources, awards were offered as monetary support or in-kind contributions. Industry partners also sponsored supplementary recognitions, including individual achievement awards. This structure celebrated both team performance and individual talent while reinforcing collaboration between academia and industry.

3.4 Implementation Plan

The hackathon was implemented over two consecutive days, with a structured agenda that balanced instruction, practice, and project-based learning. Table 2 presents the detailed timeline of activities, aligned with the intended learning outcomes (LO1–LO6).

3.5 Hackathon Details

Building on the defined learning outcomes and instructional resources, the hackathon tasked each team with developing a chatbot-based question-answering system using either a single-agent or multi-agent architecture. The system was required to (i) interact naturally with users, (ii) handle realistic situations flexibly, and (iii) deliver accurate and efficient solutions. Each prototype needed to demonstrate potential impact in *education* or *healthcare*.

Support Materials. Teams were supported by the tutorial handbook and open-source codebase, enabling them to translate conceptual knowledge into a functional chatbot, with mentors providing technical guidance as needed.

Deliverables. Final submissions required: (i) a functional prototype, (ii) presentation slides (7 minutes), (iii) a demo video (under 3 minutes), and (iv) supplementary datasets, images, or documentation for any custom-built tools.

Assessment. Team projects were assessed by an expert panel (three invited speakers and an industry representative) based on functionality, design, evaluation, innovation, and presentation quality. Mentors additionally monitored individual contributions. The detailed scoring rubric is provided in the supplementary material.

4 Results and Evaluation

4.1 Learning Materials Evaluation

To evaluate the effectiveness of our teaching resources, we adopt Tomlinson's framework (Tomlinson 2003), which emphasizes multi-dimensional rather than single-dimension evaluation. Adapted for AI agent education at the high school level, it comprises five dimensions: *Adaptability* (A), flexibility across teaching approaches; *Localization* (L), applicability in diverse contexts; *Visualization and Practice* (V), integration of visual aids and hands-on activities; *Clarity and Coverage* (C), precision and comprehensiveness of content; and *Age-Suitability* (S), appropriateness for secondary learners (Nunan 1991; Ellis 1997). The complete rubric is provided in the supplementary material. Using this rubric, we compared three representative online courses and three real-world hackathons, as summarized in Table 3. Coursera and Udemy score strongly in Clarity but fall short in Age-Suitability and Visualization, though Udemy's *AI Agents For All!* offers a no-code, beginner-friendly design that lowers entry barriers. Hackathons offer authentic practice but lack Adaptability and remain largely inaccessible to most high school students. In contrast, our hackathon shows balanced performance across all five dimensions, particularly in Localization and Age-Suitability, while maintaining strong Clarity, thereby bridging the gap between theory-focused courses and professional-level competitions.

4.2 Competition Analysis

Table 4 compares our hackathon with a range of selected international competitions. High-school contests such as ISEF, IOAI, and WAICY[15] are broad in scope and span longer periods, providing substantial preparation time but limited exposure to agentic AI. University and industry hackathons,

[12]https://www.coursera.org/projects/building-agentic-rag-with-llamaindex

[13]https://www.udemy.com/course/ai-automation-build-llm-apps-ai-agents-with-n8n-apis/

[14]https://www.udemy.com/course/no-code-ai-agents/

[15]https://www.waicy.org/

Time	Activity	Description	Linked LOs
Day 1			
07:30–08:30	Ice-breaking and team formation	Participants meet teammates, set expectations, and engage in warm-up activities.	LO6
08:30–09:15	Lecture 1: From AI to Agentic AI	Expert-led introduction covering core concepts and glossary.	LO1–LO2
09:15–09:30	Q&A session	Interactive clarification with experts.	LO2
09:30–10:15	Case study + live demo	Walkthrough and demonstration of an operational enrollment consultation chatbot.	LO2
10:15–10:30	Q&A session	Expert responses and clarifications.	LO2
10:30–11:30	Breakout discussion	Mentor-guided small-group consolidation of lecture and demo takeaways.	LO2
11:30–13:30	Lunch break	—	—
13:30–14:15	Lecture 2: Multi-Agent Systems	Architectures, workflows, protocols, and design trade-offs.	LO2
14:15–14:30	Q&A session	Interactive clarification with experts.	LO2
14:30–15:15	Guided tutorial + mini challenges	Hands-on practice with mentor support; implement basic agentic pipelines.	LO3
15:15–15:30	Break	—	—
15:30–17:00	Hackathon challenge announcement + team brainstorming	Problem statement released; teams generate draft solutions.	LO3–LO4
17:00–18:00	Expert feedback on draft ideas	Experts provide critique and improvement suggestions.	LO4–LO5
Open	Initial implementation kickoff	Teams start coding/designing solutions with mentor check-ins.	LO3–LO5
Day 2			
07:30–12:00	Hackathon sprint (team work)	Teams continue building and refining their agentic solutions.	LO3–LO6
12:00–13:00	Lunch break	—	—
13:00–13:30	Peer review exchange	Teams review each other's partial solutions and give constructive feedback.	LO5
13:30–15:30	Hackathon sprint (continued)	Final implementation and preparation of presentations.	LO3–LO6
15:30–17:00	Project presentations	Teams present prototypes to expert and industry panel.	LO5–LO6
17:00–17:30	Expert evaluation and award ceremony	Panel deliberation, scoring, announcement of winners.	LO5–LO6

Table 2: Two-day agenda of the AI Agent Hackathon with mapped LOs.

Material / Event	A	L	V	C	S
Building Agentic RAG with LlamaIndex[12]	5	3	2	4	1
AI Automation with n8n & APIs[13]	5	3	2	4	2
AI Agents For All![14]	4	4	5	3	3
Microsoft AI Agents Hackathon	3	3	3	4	2
AI Agents Hack (LabLab + MindsDB)	2	3	3	3	2
NVIDIA Agent Toolkit Hackathon	2	3	2	3	2
Our Hackathon	**5**	**5**	**5**	**4**	**4**

Table 3: Comparison of AI agent learning materials and hackathon events based on the proposed evaluation rubric.

including HackNYU[16], NASA Space Apps[17], and the Microsoft AI Agents Hackathon[18], deliver high-intensity, technically demanding experiences yet remain largely inaccessible to high school learners. Academic competitions such as the XAI Challenge target advanced participants and further widen this gap. In contrast, our hackathon adopts the 36-hour format of major events while centering on AI agent applications, supported by trained undergraduate mentors and an accessible curriculum. This design balances rigor with approachability, offering high school students a rare opportunity to engage with state-of-the-art agentic AI challenges.

[16]https://hacknyu.org/
[17]https://www.spaceappschallenge.org/
[18]https://microsoft.github.io/AI_Agents_Hackathon/

4.3 Student Learning Outcomes and Evaluation

To contextualize learning gains and assess program impact, we compared three groups: (i) *participants* who completed all instructional and hackathon activities, (ii) *controls* from the same schools who did not join, and (iii) *non-participants* from other schools who completed only the pretest. Full demographics are provided in the supplementary material.

Table 5 shows broadly comparable pretest performance across groups, with only minor variation among LOs. Non-participants scored slightly higher in LO3 and LO5, suggesting modestly stronger prior exposure, while participants and controls exhibited nearly identical profiles, validating the use of controls as a baseline.

Table 6 reports post-program performance for participants, disaggregated by grade and school type. Students showed clear improvement across most LOs, with the largest gains in LO2 (+1.02) and LO6 (+0.60). *Specialized-school* (Spec.) students scored higher than *Regular-school* (Reg.) students overall, reflecting stronger academic preparation, yet both groups displayed similar improvement patterns. Although LO4 and LO5 show slight declines, the overall trend still indicates meaningful gains in conceptual understanding, practical skills, and creative application across the cohort.

Competition	Content	Time	Mentors	Place
ISEF	STEM	Over months	Professionals	Hybrid
IOAI	AI	Several days	AI Experts	In-person
WAICY	AI	Several days	-	Hybrid
HackNYU	Beyond STEM	48 hours	Professionals	In-person
NASA Space Apps	STEM	48 hours	Professionals	In-person
XAI Challenge (IJCNN 2025) (Nguyen et al. 2025b)	LLMs	3 months	-	Online
Riding on the Back of a Whale (Nguyen et al. 2025a)	LLMs only	36 hours	Undergraduates	Hybrid
Microsoft AI Agents Hackathon	AI Agents	3 weeks	Microsoft engineers	Hybrid
AI Agents Hack (LabLab + MindsDB)	AI Agents	48 hours	Volunteers & Startup mentors	Hybrid
NVIDIA Agent Toolkit Hackathon	AI Agents	48 hours	NVIDIA professionals	Hybrid
Ours	**AI Agent Application**	**36 hours**	**Well-selected undergraduates**	**In-person**

Table 4: Comparison between our hackathon and international AI-related competitions and hackathons.

Group	LO1	LO2	LO3	LO4	LO5	LO6
Participants	2.67 / 1.83	3.12 / 1.60	3.93 / 1.47	2.97 / 1.35	3.00 / 1.55	2.02 / 1.45
Controls	2.93 / 1.83	3.01 / 1.65	3.84 / 1.56	2.74 / 1.36	3.09 / 1.60	2.20 / 1.50
Non-participants	2.58 / 1.88	2.96 / 1.70	4.18 / 1.31	3.19 / 1.24	3.57 / 1.35	2.38 / 1.43

Table 5: Pretest performance (mean / standard deviation) across groups on all LOs.

LO	Mean	Grade			School Type		Change
		10	11	12	Spec.	Reg.	
1	2.91	2.64	2.69	3.08	3.93	2.83	+0.24
2	4.15	4.04	3.97	4.33	4.47	4.12	+1.02
3	4.10	3.76	3.91	4.41	4.40	4.07	+0.16
4	2.76	2.60	2.64	2.89	3.33	2.72	-0.21
5	2.87	3.04	2.73	2.96	3.80	2.80	-0.12
6	2.61	2.88	2.31	2.73	3.00	2.59	+0.60

Table 6: Post-program scores for each LO, disaggregated by grade and school type, with pre–post score change.

4.4 Discussion

To understand how background and process factors influenced students' learning and hackathon outcomes, we examined five *Research Questions* (RQs) using Welch's t-test (Welch 1947) for group comparisons and Spearman's rank-order correlation (Spearman 1904) for non-parametric relationships. Full survey items and statistical details are reported in the supplementary material.

RQ1: Does grade level affect students' confidence and programming knowledge? Across Grades 10–12, no significant differences were found in confidence or self-rated programming skills. This indicates broadly comparable baseline readiness and supports using a unified framework for all high-school levels.

RQ2: Are there systematic confidence differences between male and female students? Male students reported significantly higher confidence and self-rated programming skills, although performance gaps were small. This mirrors known STEM self-efficacy disparities and highlights the importance of providing additional support for female learners.

RQ3: Which background and process factors are associated with hackathon performance? Prize-winning and non-winning teams showed no significant differences in motivation, prior AI knowledge, or teamwork, suggesting that scaffolding helped equalize starting points. However, students who interacted more with mentors reported higher project confidence, highlighting mentoring's role.

RQ4: How do social support and peer collaboration relate to students' motivation and learning gains? Family-and-friend encouragement correlated positively with students' interest in continuing AI learning, and students who helped classmates more often achieved higher post-test scores. Social support and peer explanation therefore appear to reinforce both motivation and conceptual understanding.

RQ5: How are AI-tool usage patterns related to conceptual understanding, and does the program change these patterns? Everyday AI use showed only a weak, non-significant link to conceptual knowledge, suggesting that frequent tool use alone does not ensure deeper understanding. Nonetheless, AI-tool usage increased significantly after the program, indicating that structured exposure promotes more confident and purposeful use.

5 Conclusion

This paper introduced a hackathon-based framework that enables high school students to explore agentic AI through scaffolded lectures, tutorials, and mentor-guided projects. The approach strengthened students' conceptual understanding, practical skills, creativity, and ethical awareness, while offering a more accessible pathway than university courses or industry hackathons. Future work will involve high school teachers more directly, adapt the framework to varied school contexts, and study longer-term effects on students' learning, ethical reasoning, and career trajectories.

References

Alonso, J. M. 2020. Teaching Explainable Artificial Intelligence to High School Students. *International Journal of Computational Intelligence Systems*, 13(1): 974–987.

Anderson, L.; and Krathwohl, D. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman. ISBN 9780801319037.

Araujo, A. A.; Kalinowski, M.; and Baldassarre, M. T. 2024. Can participation in a hackathon impact the motivation of software engineering students? A preliminary case study analysis. In *Proceedings of the 2024 IEEE/ACM 17th International Conference on Cooperative and Human Aspects of Software Engineering*, CHASE '24, 103–108. New York, NY, USA: Association for Computing Machinery. ISBN 9798400705335.

Bloom, B. S. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. London: Longmans, Green.

Broll, B.; and Grover, S. 2024. Beyond Black-Boxes: Teaching Complex Machine Learning Ideas through Scaffolded Interactive Activities. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(13): 15990–15998.

Ellis, R. 1997. The empirical evaluation of language teaching materials. *ELT Journal*, 51(1): 36–42.

Grover, S. 2024. Teaching AI to K-12 Learners: Lessons, Issues, and Guidance. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2024, 422–428. New York, NY, USA: Association for Computing Machinery. ISBN 9798400704239.

Lim, H.; Min, W.; Vandenberg, J.; Cateté, V.; and Mott, B. 2024. Unplugged K-12 AI Learning: Exploring Representation and Reasoning with a Facial Recognition Game. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(21): 23285–23293.

Long, D.; and Magerko, B. 2020. What is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, 1–16. New York, NY, USA: Association for Computing Machinery. ISBN 9781450367080.

Maldonado, D.; Cruz, E.; Abad Torres, J.; Cruz, P. J.; and Gamboa Benitez, S. d. P. 2024. Multi-Agent Systems: A Survey About Its Components, Framework and Workflow. *IEEE Access*, 12: 80950–80975.

Nandi, A.; and Mandernach, M. 2016. Hackathons as an Informal Learning Platform. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, 346–351. New York, NY, USA: Association for Computing Machinery. ISBN 9781450336857.

Nguyen, D.; Le, D.; Nguyen, L.; Vu, Q.; Le, T.; Nguyen, D.; Huynh, N.; Nguyen, H.; Tran, P.; Le, D.; Truong, S.; Koyejo, S.; Le, C.; and Quan, T. 2025a. Riding on the Back of a Whale: A Hackathon Framework for Introducing High School Students to Large Language Models. In Cristea, A. I.; Walker, E.; Lu, Y.; Santos, O. C.; and Isotani, S., eds., *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium, Blue Sky, and WideAIED*, 201–209. Cham: Springer Nature Switzerland. ISBN 978-3-031-99264-3.

Nguyen, L. S.; Vo, K. H.; Nguyen, T. H.; Bui, T. C.; Nguyen, D. Q.; Tran, T.-T.; Nguyen, A. D.; Nguyen, M. L.; Baldacci, F.; Bui, T. H.; et al. 2025b. Bridging LLMs and Symbolic Reasoning in Educational QA Systems: Insights from the XAI Challenge at IJCNN 2025.

Norouzi, N.; Chaturvedi, S.; and Rutledge, M. 2020. Lessons Learned from Teaching Machine Learning and Natural Language Processing to High School Students. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(09): 13397–13403.

Nunan, D. 1991. *Language Teaching Methodology: A Textbook for Teachers*. Upper Saddle River, NJ: Prentice Hall.

Olari, V.; Cvejoski, K.; and Eide, O. 2021. Introduction to Machine Learning with Robots and Playful Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(17): 15630–15639.

Rizvi, S.; Waite, J.; and Sentance, S. 2023. Artificial Intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. *Computers and Education: Artificial Intelligence*, 4: 100145.

Russell, S.; and Norvig, P. 2009. *Artificial Intelligence: A Modern Approach*. USA: Prentice Hall Press, 3rd edition. ISBN 0136042597.

Sajja, R.; Ramirez, C. E.; Li, Z.; Demiray, B. Z.; Sermet, Y.; and Demir, I. 2024. Integrating Generative AI in Hackathons: Opportunities, Challenges, and Educational Implications. *Big Data and Cognitive Computing*, 8(12).

Spady, W. G. 1995. Outcome-based education : critical issues and answers.

Spearman, C. 1904. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1): 72–101.

Surendran, S.; Mack, K.; Bingham, N. M.; Edwards, N.; Frost-Schenk, J.; Keshishi, N.; Matos, F.; Moldoveanu, J.; Walsha, R.; and Bodman-Smith, K. 2023. The use of extracurricular hackathons to promote and enhance students' academic and employability skills. *International Journal of Educational Research Open*, 5: 100307.

Tan, X.; Cheng, G.; and Ling, M. H. 2025. Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8: 100355.

Tomlinson, B. 2003. Developing principled frameworks for materials development. In Tomlinson, B., ed., *Developing Materials for Language Teaching*, 107–129. London: Continuum.

Touretzky, D.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019. Envisioning AI for K-12: what should every child know about AI? In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference* *and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI'19/IAAI'19/EAAI'19. AAAI Press. ISBN 978-1-57735-809-1.

Varuvel Dennison, D.; Garcia, R. C. C.; Sarin, P.; Wolf, J.; Bywater, C.; Xie, B.; and Lee, V. R. 2024. From Consumers to Critical Users: Prompty, an AI Literacy Tool for High School Students. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(21): 23300–23308.

Wang, S.; Wang, F.; Zhu, Z.; Wang, J.; Tran, T.; and Du, Z. 2024. Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252: 124167.

Welch, B. L. 1947. The generalization of "Student's" problem when several different population variances are involved. *Biometrika*, 34(1-2): 28–35.

Wiese, L. J.; Patil, I.; Schiff, D. S.; and Magana, A. J. 2025. AI ethics education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 8: 100405.

Yim, I. H. Y.; and Su, J. 2025. Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education*, 12(1): 93–131.

Zhang, D.; Wu, X.-J.; and Yu, J. 2021. Label Consistent Flexible Matrix Factorization Hashing for Efficient Cross-modal Retrieval. *ACM Trans. Multimedia Comput. Commun. Appl.*, 17(3).