AI-driven NPCs with Dynamic Storylines for an Educational Game

[1] Ananya P S, [2] Shirisha Reddy V, [3] Yeshas R
[1][2][3] Presidency University, Bengaluru, India
[1] ananyakalluraya@gmail.com , [2] shirishareddyv687@gmail.com , [3] yashassonu1222@gmail.com

Abstract— ***While educational games have significant potential, right now this field currently struggles with many fundamental limitations. They often use the same fixed, static, pre-scripted narratives that fail to encourage student engagement and personalized learning. We address this gap by introducing a new approach to educational game design, using Large Language Models. This approach goes beyond just following a script and instead creates a game that changes and adapts as you play. Our system's key innovation lies in providing non-player characters (NPCs) intelligent, historically inspired personalities, to act as responsive mentors rather than merely repeating out preset dialogues. These NPCs keep track of each student's knowledge profile, progress, and learning goals;the game's story and challenges change real time based on the user. This means the characters talk and work together in ways that feel natural and meaningful, helping students explore different subjects in a deeper way. This real time adaptation orchestrates dynamic dialogues and cross-domain collaborations among NPCs, facilitating authentic interactions. Our goal is to make learning more exciting, help students remember what they learn better, give them a more lasting understanding of concepts than traditional, scripted educational tools can offer.***

Index Terms— Adaptive Learning, AI Planning, Cross-Domain Collaboration, Dynamic Narratives, Educational Games, Generative AI, Generative Agents, Human-Computer Interaction, Large Language Models (LLMs), Multi-agent Systems, Non-Player Characters (NPCs), Personalized Learning, Persona-Driven AI, Scaffolding, Student Engagement, Unity Game Engine.

INTRODUCTION

Traditional educational games lack engaging, personalized, and adaptive learning experiences, frequently resulting in reduced knowledge retention and waning curiosity among students [1]-[3]. Many games rely on linear, pre-scripted narratives that prevent dynamic interaction and limit deep, curiosity-driven learning, which can hinder knowledge retention and further exploration of topics beyond the core curriculum [1][3][4]. Research shows that students often find traditional educational games and passive learning approaches disengaging, as these methods are not tailored to individual interests or needs [1][5][6]. A systematic review of adaptive game-based learning found that games that do not personalize experiences or respond to individual learner profiles lead to decreased engagement and lower motivation compared to adaptive alternatives, which better support student curiosity, needs, and performance [6]. Furthermore, evidence across several research studies, including findings from focus groups, surveys, and case analyses, indicates that while conventional educational games can have some positive impacts, students and educators overwhelmingly favour features such as customization, adaptability, and immediate feedback – the design elements absent in most traditional games [1]-[6].

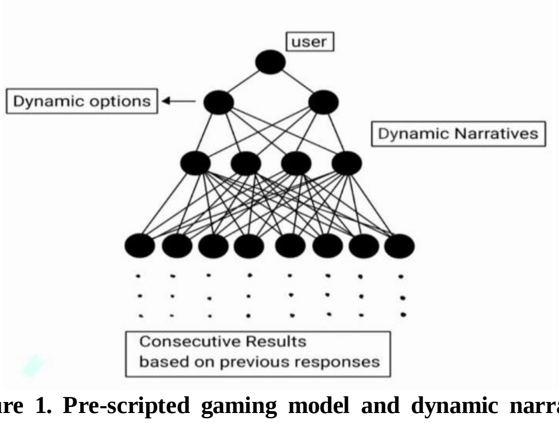

Figure 1. Pre-scripted gaming model and dynamic narrative model

Prior research demonstrates that adaptive educational games "performed better for the students" in terms of engagement and learning performance, as shown in studies comparing adaptive and non-adaptive game environments [6, p. 5]. Emphasizing the value of personalization, this research notes that adjusting content and support to match individual learners is essential for sustaining interest and promoting deeper learning. A recent systematic review reinforces this view, stating that "adapting the content and the experience to each learner usually leads to better user retention and a more meaningful and deeper learning experience" [6, p. 1].

Beyond the limitations of stagnant educational games, similar challenges emerge in the use of chatbots for learning environments. While these systems are often introduced as supportive tools, chatbots without defined personas or visual characters tend to lack a human touch, making interactions feel impersonal and transactional. This absence of relatability can cause students to disengage, with many describing text-only chatbots as akin to "talking to a wall." [7, p. 7] In contrast, research highlights that persona-driven or avatar-supported interfaces create more engaging and meaningful interactions, helping learners feel more connected to the system and sustaining their motivation over time [7].

OBJECTIVES

At its heart, this research work is about changing how we learn. We want to move away from passive learning of information and towards an active learning, where the story itself is the teacher. We believe games are the perfect way to do this as they can be tailored to each person and provide instant feedback [3][7].

To bring this vision to life, our research work has a few key goals. We are:

1. Building a working game prototype in Unity that includes smart AI characters (NPCs) that can hold dynamic conversations [3][7][8].
2. Giving each character a unique personality and expertise. You might meet Professor Stein, a science expert, or Sir William, a literature guide. What's unique is that these characters can actually work together. If a player asks a question that touches on multiple subjects, the NPCs can consult one another, creating a truly collaborative learning experience [9][10].

3. Using a smart AI system that blends a creative language model which is for natural dialogue with a logical planner, this in turn allows the game's story to adapt on the fly, creating challenges and support that are perfectly matched to each student's progress and needs [1][3][11].

Finally, we want to show that generative AI can create learning tools that are not just more engaging, but also genuinely effective.

Related Work

AI Planning and Dynamic Storylines

For years, creating truly dynamic stories in games has been a very big challenge. In early stages, games relied on simple, rule-based systems, where stories were often rigid, predictable, and couldn't easily grow when things got complicated [9][12].

Development came with better AI planning methods like Hierarchical Task Network (HTN) planning. This was a perfect match for storytelling because it breaks down big and complex goals into smaller and manageable steps which is much like how a main quest in a game is divided into smaller side-quests and objectives [9][10][12].

Today, even more powerful approaches are available. Newer models like GOALNET are more efficient and robust [10][12]. The real breakthrough, however, comes from using large language models (LLMs) in systems where multiple AI agents can work together. These advanced agents can process more than just text. They can integrate images and video thus allowing them to perceive and reason in ways that create much better realistic simulations of the real world [9].

What all this technology means is that we can finally create game characters that are no longer puppets following a script. They can understand the nuances of human language, react to unexpected situations, and adapt their conversations and actions in real-time. This allows them to guide players on a truly personalized learning journey, even when the player does something completely unforeseen [9]-[11].

AI Framework: A Hybrid Approach

The AI framework combines a large language model (LLM), symbolic planning, and multi-agent systems to create an adaptive and intelligent learning environment [9]. The LLM acts as the main "brain," handling natural language input and generating meaningful responses [10]. Planning models like HTN and GOALNET break down big goals into smaller tasks, ensuring that lessons follow a clear and structured path [12]. At the same time, multi-agent protocols allow different AI tutors or NPCs to work together, sharing expertise when one agent gets "stuck," much like real classroom collaboration [13]. Built-in error management, such as detecting hallucinations or keeping conversations grounded, helps maintain accuracy and trust [14]. This means the system feels more responsive, reliable, and engaging. NPCs can adapt lessons to individual needs, use natural and immersive dialogue, and create a more interactive and motivating learning experience [1][3].

Generative agents simulate believable human behaviour by dynamically managing extensive memory streams, reflection, and planning modules [10][14]. They experience, remember, and reflect on interactions, coordinating social behaviours autonomously in a game environment akin to The Sims, producing emergent social dynamics [5][12]. This architecture integrates a retrieval-augmented memory mechanism for long-term coherence, enabling NPCs to plan, act, and replan dynamically, react contextually, and evolve their relationships [2][7]. Solo Performance Prompting (SPP) enables a single LLM model to act as multiple personas collaborating internally, enhancing cognitive synergy for improved problem solving, creativity, and reduced hallucinations [2][15]. Personas are dynamically identified per task context, facilitating brainstorming, critique, and iterative refining of outputs, resulting in more accurate and nuanced NPC dialogues [2][10].

The Role of AI Personas and Avatar Design

Incorporating players' actions, environmental state, and intent into LLM NPC dialogue generation creates more immersive, relevant, and adaptive conversations [16]. For example, NPCs recognizing objects a player points to or responding to dynamic game states increase player satisfaction, engagement, and perceived realism in RPGs [3]. The design of NPC avatars influences student engagement and learning experiences in AI tutor platforms. Realistic avatars tied to known instructors enhance credibility and trust, while diverse avatar types accommodate different learner preferences, affecting the style and scope of learner-NPC interactions [4][7].

There's a real conflict with AI tutors that seem human. On one hand, their human-like qualities can make learning very much engaging and fun. On the other hand, these same human-like qualities can become a distraction that makes it hard for a student to focus on the lesson [4][7]. One study found that when AI characters in a learning game were designed to give smart, relevant answers, students were more engaged and felt the AI was genuinely useful [5]. But you have to be very careful, if an AI tutor is too personable or friendly, it can lead to distraction among the learners which might lead to pulling of attention away from the actual goal of learning [7].

The main challenge for this research work is telling a story that's fun but still teaches the lesson. To address this, we tailor each NPC's persona to the specific learning context [10]. For example, in a science task, students interact with Professor Stein. On the other hand, a creative writing exercise is guided by Sir William, whose playful persona appeals to "engagement-oriented" students who learn best through fun, interactive characters [6][10]. Using these subject-specific personas helps ensure that the guidance and feedback match both the topic and the learner's style [2]. By building unique, memorable characters, the system delivers an immersive experience without losing sight of its main educational purpose.

The Pedagogical Framework: Scaffolding with Generative AI

It's clear that using advanced AI is now important for teaching. Games are now powerful tools because they give every student a personalized experience with instant feedback [4]. The great thing is, this technology isn't just for one subject, it's flexible enough to help students learn everything including writing, science and math [4]. In our research work, we trust that AI-driven story itself is the main teaching tool. The game doesn't depend on traditional lessons. Instead, the changing plot and the smart characters are the main way content and support are delivered. The AI constantly checks on how a student is doing and adapts the story accordingly, it might introduce a new challenge, offer a hint, or adjust the difficulty to the perfect level to keep the student engaged.

PROPOSED SYSTEM ARCHITECTURE AND DESIGN

The proposed system will be built on a robust, multi-component technology stack, with the large language model at its core.

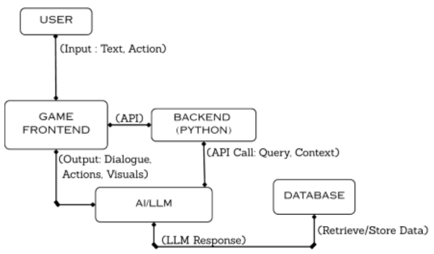

Figure 2. System Architecture design.

The system uses a modular, three-tier design [1] that allows for real-time, context-aware interactions powered by a Large Language Model (LLM). This layout, shown in Figure 2, keeps different parts of the system separate so they can work efficiently on their own tasks.

Presentation Tier: The Game Frontend

As shown in Figure 2, the Game Frontend is the main way users interact with the system. Its functions include capturing user input, which can be either text (like natural language questions or conversation) or specific actions (like predefined game commands). It then sends the performed input to the application layer through a standardized API call. It is also responsible for displaying the system's response, which includes generated dialogue, available actions (like buttons or menus), and updated visuals showing the current environment or story status.

Application Logic Tier: The Backend (Python)

The Backend is built using Python and acts as the central control unit of the system. It receives requests from the Frontend via an API and has important duties like Handling and validating the initial request from the Frontend and keeping track of the ongoing interaction, including the conversation history, game variables, and environmental details [2]. It creates a structured request with the user's input and the current context, then sends it to the AI/LLM via an API. This ensures the LLM gets only the needed and relevant information to create a logical response [16].

Intelligence and Data Tier: AI/LLM and Database

This part of the system includes the intelligence and data storage needed for it to function properly (see Figure 2). Large Language Model (LLM) is the main part that creates dynamic content for the system using the context and query from the Backend. It connects with the Database to fetch static info like character traits or conversation details, or to save dynamic data like progression or user preferences that affect future responses [2][10][16]. It processes the input and data to generate the final response, which is then sent back to the Frontend. Database is the storage part that keeps all the necessary data for the system's operation, ensuring that the system stays consistent across different sessions.

Data Flow and Communication Protocol

The system works using a clear request-response process, User input moves from the Frontend to the Backend, which then sends a request to the AI/LLM. The AI/LLM uses the Database to get the needed info and then sends the generated response back to the Frontend for display. As shown in Figure 1, all communication between parts of the system uses standardized API protocols, making it easy to scale and maintain. This setup keeps the heavy work of the LLM separate from the user interface and application logic layers [1].

Persona-Driven NPCs

In a multi-agent system, each character is a generative agent, a computational software agent that simulates believable human behaviour. The behaviour of these agents is conditioned by their past experiences and current environment [9]. These agents are designed with distinct personalities, goals, and beliefs, much like characters in a good story [9][10][12]. For example, a generative agent's personality can be defined by a set of traits and a profile that describes its role in natural language [10][12]. In our proposed system, their dialogue, tone, and knowledge delivery are consistent with their unique personas, which are inspired by historic figures. For example, our English mentor that is Sir William, speaks like a dramatic playwright. On the other hand, the science mentor, Professor Stein, explains things with the logic of an eccentric physicist (refer Figure 4). This makes learning feel like a real conversation that we have with an expert, which helps students engage more deeply and remember better [2][7].

But we also had to find the right balance. A character that's too social can become a distraction. That's why our system can adjust an NPC's personality, making sure its chattiness doesn't get in the way of learning [7]. This gives learners control over the interaction and keeps the experience both enjoyable and efficient [5][7].

Cross-Domain Collaboration

Our system's key stand out feature is its backend logic, which uses a multi-agent framework which is connected to a large language model (LLM) to manage NPC interactions [9]. This approach allows decentralized, autonomous agents to work together to achieve collective goals [9][11]. The system is designed for NPCs to cooperate by exchanging state information, assigning responsibilities, and coordinating actions through established communication protocols [9][11]. When an agent's knowledge is insufficient to fully answer a player's query, the system identifies this and performs a "handoff" to a more suitable character [2][9][11]. This is because agents can access a variety of available tools, including other agents, to gather missing information when they lack the full knowledge needed for a subtask [9]. This mimics a real-world expert network, not only enhancing the learning experience but also adding a unique layer of narrative depth to the game. Multi-agent systems are a more accurate representation of the real world, as many real-world applications naturally involve multiple decision-makers interacting simultaneously [9].

Dynamic Storyline and Personalized Learning

A dynamic, narrative-centered game can scaffold learning by coupling adaptive NPC behaviours with the learner's evolving profile [3][6], so that conversation traces and assessed states guide an LLM-driven backend to adjust plot beats, dialogue, and supports in real time to remediate misconceptions, offer just-in-time help, and calibrate challenge difficulty to the learner's level [2][6][14]. Unlike fixed scripts or static hint tiers, this iterative adaptation blends planning, memory, and reflection in agents to retrieve salient context and personalize interactions, transforming each hint into a context-rich practice moment embedded in story progression rather than a generic prompt [5][14]. Such tailored scaffolding aligns with evidence that adaptive game-based learning boosts engagement, self-efficacy, and performance by matching task demands to competence and supplying immediate, meaningful feedback [1][2][4]. As a result, the game becomes a space where learners expect success, persist through difficulty, and develop a growth mindset—benefits repeatedly observed when narratives, adaptive supports, and timely feedback are combined in educational games and AI-mediated tutors [5][7]. This stands in contrast to traditional systems that rely on pre-authored dialog trees or fixed difficulty curves, which cannot continuously personalize challenges and assistance to the learner's moment-to-moment needs [1][4][6].

The research work will follow a phase-based approach to design, implement, and validate the proposed adaptive educational game. Initially, a detailed literature review will inform system architecture design, character persona planning, and knowledge base structuring for NPCs [1][2][3][6][9]. The core implementation phase includes developing a Unity prototype [8], integrating LLM APIs for dynamic dialogue [11], and implementing cross-domain collaboration logic for NPC interactions [9][10][11], followed by unit and integration testing [14]. The final phase focuses on refining character interactions, narrative flow, user interfaces [4][7], and compiling comprehensive documentation.

To evaluate the system, a mixed-methods user study will compare student engagement and task performance in the adaptive game against a static equivalent [1][3][4][6]. Both automated evaluation (e.g., "LLM-as-judge" for free-form responses [2][15]) and human assessment [14][16] will ensure accuracy, detect hallucinations, and identify subtle biases.

RESULTS

We have successfully created and tested a prototype of the conversational flow model for our educational game. The main focus of this stage was to ensure the AI tutors' conversational flow and their distinct personalities were working correctly.

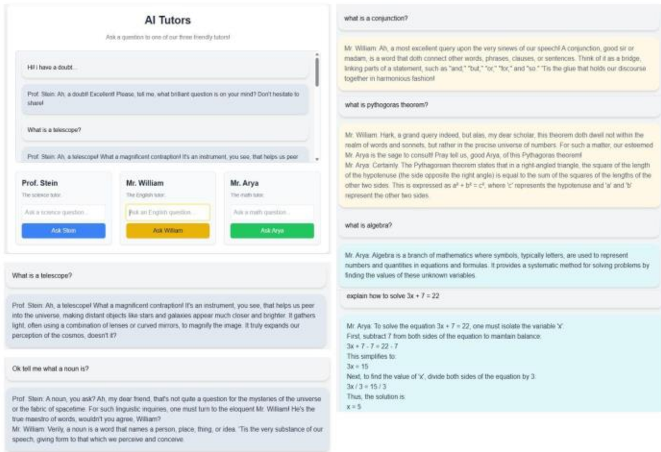

Figure 3. Conversational model design.

Our tests confirmed that the three AI tutors - Professor Stein (science), Mr. William (English), and Mr. Arya (math), each maintain a unique and consistent persona. For example, Professor Stein uses a knowledgeable, eccentric tone, while Mr. William speaks with poetic flair. Mr. Arya, on the other hand, provides direct and systematic explanations (refer Figure 3).

A key achievement is the cross-domain collaboration feature. When a query is outside of an AI's expertise, the system successfully "hands off" the conversation to the correct tutor. For instance, when asked about the Pythagorean theorem, Mr. William correctly deferred to Mr. Arya, who then provided the accurate mathematical explanation (refer Figure 3).

Figure 4: 3D model design and structure.

The current system architecture is functional, with a Unity-based [8] frontend communicating with a Python backend that uses an AI/LLM API. This setup allows for dynamic dialogue generation and lays the groundwork for saving conversation data, which is essential for future personalized learning features.

Refer Figure 3, the Unity [8] Persona Designs include (refer Figure 4):

- Professor Stein: The science tutor. He is an eccentric physicist who conveys complex principles with a calm, rational tone.
- Mr. William: The English tutor. He is a playwright who speaks with flair and poetic language.
- Mr. Arya: The math tutor. He provides direct and systematic explanations of mathematical concepts.

CHALLENGES

AI systems face several significant challenges in educational and interactive contexts. Long-term memory retention and context handling remain major technical hurdles. As noted in the persona research, "incorporating user history data into the prompt for personalizing LLMs could lead to input exceeding context length as well as increased inference costs" [17, p. 8]. In group interactions, determining when an AI agent should contribute is non-trivial: research reveals that "participants raised significant concerns about feeling overwhelmed and interrupted by the proactive variant" [13, p. 2] of AI agents, highlighting the complexity of appropriate response timing. Educational applications must balance individualized learning with standardized curricula. The research indicates that "traditional learning platforms, such as classroom training, follow an old technique of teaching called the 'one-size-fits-all' approach" [1, p. 2], but personalizing AI-driven instruction presents implementation challenges. Finally, bias and safety concerns emerge when assigning personas or roles to AI agents. The persona research shows that "assigning personas to LLMs aids in jailbreaking" and that "LLMs consistently exhibit toxicity in a range of topics when assigned personas." Additionally, "applying role-play often increases the overall likelihood of generating stereotypical and harmful outputs" [17, p. 8]. These challenges collectively illustrate the technical, social, and ethical complexity of deploying AI agents in dynamic, educational, and socially interactive settings.

CONCLUSION

This proposal outlines the design and development of a novel educational game featuring LLM-driven NPCs and dynamic storylines, offering a deeply personalized and adaptive learning experience that overcomes the rigidity of traditional educational software. The system is designed to enhance student engagement and motivation by providing a compelling, goal-oriented narrative that dynamically adjusts to the learner's level, helping students persist with challenging topics and achieve a sense of success. This tailored scaffolding is projected to improve task performance and knowledge retention compared to static educational media.

The research work novelty lies in its strategic integration of advanced AI methodologies to address a complex, interdisciplinary problem. Beyond simple generative AI tools, the system leverages LLMs to create a structured, interactive narrative guided by a dynamic sociability model, which balances student enjoyment with effective learning. A key innovation is the cross-domain collaboration model, which enables students to explore multiple topics and understand how different areas of knowledge intersect, guided by an intelligent network of adaptive virtual experts. Overall, the proposal represents a significant step forward in applying artificial intelligence to create transformative, engaging, and effective educational tools.

FUTURE WORK

Future research will explore long-term impacts, including metacognitive skill development and integrating the "Learning By Teaching" paradigm, where students explain concepts back to NPCs to enhance knowledge retention [5][6]. This roadmap lays the foundation for developing intelligent, adaptive, and engaging learning companions for diverse learners [1][3][4].

REFERENCES

[1] R. Raguraman, S. Pasupathi, and J. Raju, "Adaptive NPC in Serious Games Using Artificial Intelligence," SSRN, Jan. 2024. doi: 10.2139/ssrn.4806061.

[2] Y. Liu, P. Sharma, M. Oswal, H. Xia, and Y. Huang, "PersonaFlow: Boosting Research Ideation with LLM-Simulated Expert Personas," arXiv preprint arXiv:2409.12538, 2024. doi: 10.48550/arXiv.2409.12538.

[3] L. Matyas and Csepregi, "The Effect of Context-aware LLM-based NPC Conversations on Player Engagement in Role-playing Video Games," in Proceedings of a Conference [Online]. Available: https://api.semanticscholar.org/CorpusID:268859294.

[4] A. Evmenova, K. Regan, R. Mergen, and R. Hrisseh, "Educational Games and the Potential of AI to Transform Writing Across the Curriculum," Education Sciences, vol. 15, no. 5, p. 567, May 2025. doi: 10.3390/educsci15050567.

[5] W. D. W. Gonzales, D. Shen, A. Yan, N. Xie, M. Francisco, and P. Wong, "AI NPCs in an Educational Metaverse: Evaluating the Effectiveness of Prompt Templates for Contextual Interactions," in Artificial Intelligence in Education and Learning Analytics, Singapore: Springer, 2025, pp. 53–74. doi: 10.1007/978-981-96-4952-5_4.

[6] D. Chiotaki, V. Poulopoulos, and K. Karpouzis, "Adaptive game-based learning in education: a systematic review," Frontiers in Computer Science, vol. 5, Jun. 2023. doi: 10.3389/fcomp.2023.1062350.

[7] C. T. Tan, I. Atmosukarto, B. Tandianus, S. Shen, and S. Wong, "Exploring the Impact of Avatar Representations in AI Chatbot Tutors on Learning Experiences," in Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '25), 2025, pp. 1–12. doi: 10.1145/3706598.3713456.

[8] Unity Technologies, "Unity Real-Time Development Platform | 3D, 2D, VR & AR Engine," Unity, accessed Sep. 14, 2025. [Online]. Available: unity.com.

[9] S. Chen, Y. Liu, W. Han, W. Zhang, and T. Liu, "A Survey on LLM-based Multi-Agent System: Recent Advances and New Frontiers in Application," arXiv preprint arXiv:2412.17481, 2025. [Online]. Available: https://arxiv.org/abs/2412.17481.

[10] J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative Agents: Interactive Simulacra of Human Behavior," arXiv preprint arXiv:2304.03442, 2023. [Online]. Available: https://arxiv.org/abs/2304.03442.

[11] S. Huang, N. Lipovetzky, and T. Cohn, "Planning in the Dark: LLM-Symbolic Planning Pipeline without Experts," arXiv preprint arXiv:2409.15915, 2024. [Online]. Available: https://arxiv.org/abs/2409.15915.

[12] S. Cooper, A. El Rhalibi, M. Merabti, and M. Price, "Dynamic Interactive Storytelling for Computer Games Using AI Techniques," Jan. 2010.

[13] S. Houde, K. Brimijoin, M. Muller, S. I. Ross, D. A. S. Moran, G. E. Gonzalez, S. Kunde, M. A. Foreman, and J. D. Weisz, "Controlling AI Agent Participation in Group Conversations: A Human-Centered Approach," arXiv preprint arXiv:2501.17258, Jan. 2025. [Online]. Available: https://arxiv.org/abs/2501.17258.

[14] Y. Wu, S. Liang, C. Zhang, Y. Wang, Y. Zhang, H. Guo, R. Tang, and Y. Liu, "From Human Memory to AI Memory: A Survey on Memory Mechanisms in the Era of LLMs," arXiv preprint arXiv:2504.15965, 2025. [Online]. Available: https://arxiv.org/abs/2504.15965.

[15] Z. Wang, S. Mao, W. Wu, T. Ge, F. Wei, and H. Ji, "Unleashing the Emergent Cognitive Synergy in Large Language Models: A Task-Solving Agent through Multi-Persona Self-Collaboration," arXiv preprint arXiv:2307.05300, 2024. [Online]. Available: https://arxiv.org/abs/2307.05300.

[16] S. Houde, K. Brimijoin, M. Muller, S. I. Ross, D. A. S. Moran, G. E. Gonzalez, S. Kunde, M. A. Foreman, and J. D. Weisz, “Controlling AI Agent Participation in Group Conversations: A Human-Centered Approach,” in Proc. 30th Int. Conf. Intelligent User Interfaces (IUI ’25), New York, NY, USA, 2025, pp. 390–408. doi: 10.1145/3708359.3712089. [Online]. Available: https://doi.org/10.1145/3708359.3712089.

[17] Y.-M. Tseng, Y.-C. Huang, T.-Y. Hsiao, W.-L. Chen, C.-W. Huang, Y. Meng, and Y.-N. Chen, “Two Tales of Persona in LLMs: A Survey of Role-Playing and Personalization,” arXiv preprint arXiv:2406.01171, 2024. [Online]. Available: https://arxiv.org/abs/2406.01171.

18. Python Software Foundation, "Welcome to Python.org," Python.org. [Online]. Available: https://www.python.org/. [Accessed: 14-Sep-2025]. [7] E. C. e. al, "Energy management and Industry 4.0: Analysis of the digital transformation of energy efficiency programs," Applied Energy, 2025.