Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 1 (2026)

Edizioni Universitarie Romane
editoria scientifica dal 1952

AI-mediated socio-emotional support for children with autism: a human-supervised framework

Lucia Campitiello [1], Anna Rescigno [1], Emanuela Zappalà [1], Amelia Lecce [2], and Stefano Di Tore [1,*1]

[1] University of Salerno; lcampitiello@unisa.it; annrescigno@unisa.it; ezappala@unisa.it; sditore@unisa.it
[2] University of Sannio; alecce@unisannio.it;
* Correspondence: lcampitiello@unisa.it

Abstract: Artificial Intelligence (AI) is progressively reshaping the field of inclusive education by offering new tools to support structured and personalized learning processes for students with autism spectrum disorder (ASD). This contribution introduces EMORI, a conceptual and architectural framework for the use of artificial intelligence as a mediated educational support in socio-emotional interventions for children with autism. The project is implemented through EmoRobot, an embodied robotic agent, and supported by Socrate, an AI orchestration platform designed to manage conversational logic through pedagogically guided retrieval-augmented generation techniques. The system is grounded in the socio-emotional intervention program proposed by Patricia Howlin and adopts a state-based conversational model that structures interaction into progressive and predictable stages. Within each state, the language model operates under constrained generative conditions, providing adaptive prompts and feedback while preserving instructional control. A central feature of the framework is the adoption of a human-in-the-loop approach, ensuring continuous pedagogical supervision. The contribution of this work lies in the definition of a theoretical and methodological model for the pedagogically sustainable integration of artificial intelligence into socio-emotional interventions for children with autism.

Keywords: Artificial Intelligence; Inclusive Education; Autism Spectrum Disorder; Socio-Emotional Learning; Human-in-the-Loop Systems.

[1]Authorship attribution:

Lucia Campitiello: Conceptualization, Methodology, Theoretical analysis, Framework development, Writing-Original Draft, Writing-review & editing.

Anna Rescigno: Literature review, Framework development, Writing-review & editing.

Emanuela Zappalà: Literature review, Writing-review & editing.

Amelia Lecce: Literature review.

Stefano Di Tore: Supervision, Project administration, Writing-review & editing

All authors have read and agreed to the published version of the manuscript.

1. Introduction

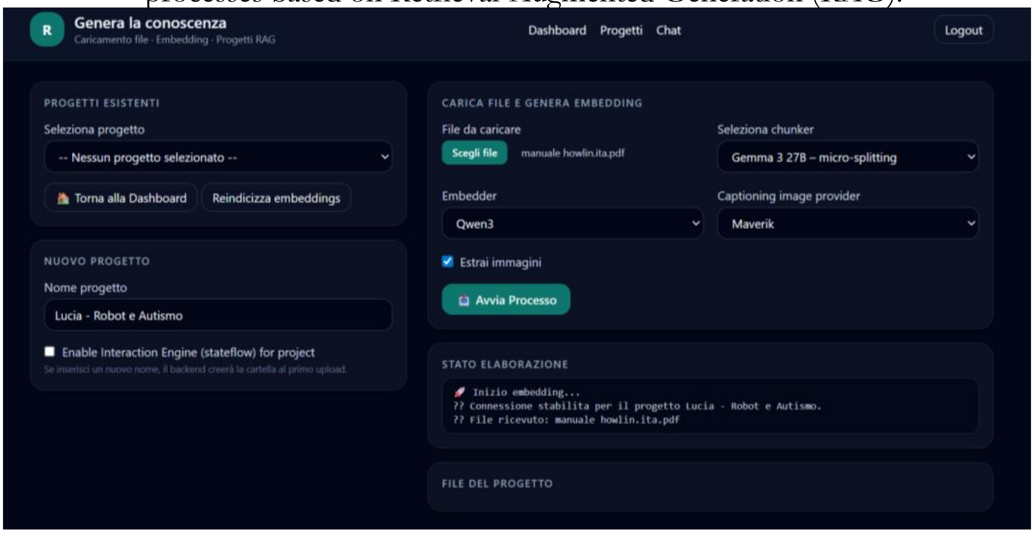
Artificial Emotional Intelligence Artificial Emotional Intelligence (EAI) has become an important research area within human-machine interaction and technology-mediated learning, focusing on the use of artificial intelligence to support, mediate, and scaffold socio-emotional learning processes. In educational contexts, EAI does not imply autonomous emotional awareness, but rather the design of structured and adaptive digital environments capable of guiding emotional understanding and interaction. From a theoretical perspective, EAI integrates contributions from cognitive psychology, neuroscience, computational linguistics, and software engineering. Foundational models such as Ekman's theory of basic emotions, appraisal-based approaches to emotion, and dimensional models of affect have informed the development of emotion-related computational systems. As noted by Floridi (2022), the integration of artificial intelligence into domains traditionally associated with human cognition and affect entails a significant epistemological shift, redefining how emotional processes are represented and mediated through technology. In the context of autism spectrum disorder (ASD), the pedagogical relevance of AI-mediated socio-emotional support becomes particularly evident. Children with autism often show atypical processing of emotional stimuli and difficulties in recognizing and interpreting others' emotions (Baron-Cohen, 1995). In response to these challenges, educational and social robotics have increasingly been explored as tools for fostering inclusion and socio-emotional learning. While educational robotics primarily targets cognitive and problem-solving skills, social robotics emphasizes relational interaction and communication through embodied agents. The convergence of these approaches within the field of Socially Assistive Robotics has shown promising results in supporting participation, joint attention, imitation, and communication in children with ASD (Scasellati et al., 2018; Rudovic et al., 2018). These effects are commonly attributed to features such as structured repetition, reduced social anxiety, personalized interaction, and immediate feedback, which align with established theories of social learning (Bandura, 1977) and Theory of Mind frameworks (Premack & Woodruff, 1978). In autism, difficulties in attributing mental states to others, conceptualized as *mindblindness* (Baron-Cohen et al., 1985), represent a key challenge for socio-emotional development. Within this theoretical landscape, the structured educational methodologies developed by Patricia Howlin constitute a foundational reference for intervention in ASD. Howlin's programme is based on the systematic application of cognitive-behavioural principles and proposes a hierarchical progression from basic emotion recognition to more complex socio-cognitive abilities, such as understanding intentions and false beliefs (Howlin et al., 1999). Building on this framework, the EMORI project can be interpreted as an AI-mediated extension of Howlin's educational principles, aimed at supporting the development of socio-emotional competencies in children with autism. The project is implemented through EmoRobot and is supported by Socrate, an artificial intelligence orchestration system. By combining structured activity sequences inspired by Howlin's programme, EMORI aims to create predictable and pedagogically grounded environments for socio-emotional learning, fostering emotional understanding and the generalization of acquired competencies to everyday life contexts.

2.The EMORI System Architecture

The EMORI project (Emotional Mediation Oriented Robotic Interaction) is embedded within a broader human–robot interaction framework, in which the pedagogical and methodological principles outlined are implemented through a specific technological architecture designed to support educational intervention. Within this framework, the educational intervention is realized through the synergistic integration of a conversational agent and an open-source robotic system, enabling action both at the software level, through artificial intelligence–supported dialogic processes, and at the embodied level, via direct interaction with a physical robot. In this configuration, the Socrate system operates as an artificial intelligence orchestration layer, managing conversational logic, state-based structured interaction progression, and adaptive feedback strategies, while EmoRobot assumes the role of an embodied social mediator supporting educational interaction.

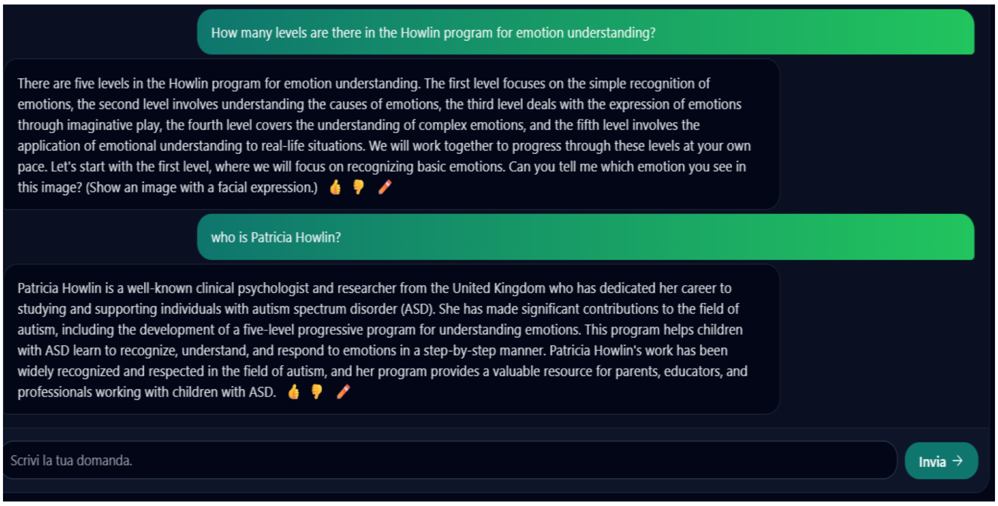
Specifically, Socrate (Orchestrated RAG-Based Conversational System for Learning, Tutoring, and Engagement) is an artificial intelligence service coordination platform developed by the TIRESIA (Interuniversity Team for Educational Research on Artificial Intelligence). The platform is designed to support the development of educational-purpose conversational agents through a modular, provider-agnostic architecture, enabling the flexible integration of heterogeneous language models and diverse computational services. This design approach allows the agent's behavior to be adapted to specific educational intervention objectives, while avoiding dependence on a single technological solution.

Figure 1: Interface of the Socrate platform for knowledge generation and embedding management, used for the orchestration of conversational flows and learning processes based on Retrieval-Augmented Generation (RAG).



Within this configuration, the Socrate platform constitutes the infrastructural layer through which the educational objectives of the EMORI project are translated into an operational and controlled conversational logic. EMORI aims to develop an educational chatbot designed to support socio-emotional learning in children with autism spectrum disorder. In particular, the system has been trained on the contents of the socio-emotional intervention program developed by Patricia Howlin (Howlin et al., 1999), focused on emotion recognition and comprehension. The activities outlined in the program have been integrated into the platform to support a guided interaction consistent with the hierarchical levels of the intervention. The EMORI system pursues the following objectives: 1) to support predictable educational interactions by incorporating the hierarchical progression of the Howlin program; 2) to enable personalized interaction through the modulation of linguistic complexity and dialogic structure in response to user performance; 3) to ensure pedagogical supervision by maintaining a human-in-the-loop approach, in which educators retain an active role in monitoring the agent's behavior.

Figure 2: Example of a structured interaction flow with predefined prompts implemented within the EMORI system



From an architectural perspective, the system implemented within the EMORI project operates according to a sequential workflow articulated into three main phases: data preparation, information retrieval, and reasoning. This configuration is consistent with contemporary approaches based on retrieval-augmented generation and cognitive agents grounded in large language models (Lewis et al., 2020). Within this framework, large language models are not employed merely as generators of textual output, but assume a role of cognitive and communicative regulation, coordinating the construction of the interactive context, the formulation of prompts, and the management of feedback mechanisms throughout the interaction. In line with this architectural approach, which integrates cognitive, communicative, and embodied dimensions of interaction, the EMORI project includes the integration of a dedicated application for facial expression recognition, used in synergy with EmoRobot to provide the child with immediate feedback on the imitation of presented emotional expressions. This functionality is conceived as an operational extension of the conversational workflow, supporting guided emotion learning through a mechanism of immediate and controlled reinforcement, consistent with an educational rather than evaluative logic. Accordingly, the application does not serve diagnostic purposes but functions as a support tool for imitative practice, providing feedback on the degree of correspondence between the produced expression and the proposed visual model.

3. Conversational design and state progression

The conversational logic of the EMORI system is grounded in a state-based interaction model, in which dialogue is organized as a finite set of discrete states connected by conditional transitions. This design choice aims to ensure a high degree of structure, predictability, and pedagogical coherence throughout the entire educational pathway, addressing the specific needs of children with autism, for whom overly open or unpredictable interactions may lead to cognitive overload or anxiety. Each state is associated with a specific pedagogical function (e.g., initial welcome, orientation, pathway selection, instruction, guided practice, and feedback), while transitions between states are triggered based on user input, user performance, and predefined criteria. In this way, the system maintains a controlled interactive environment while allowing a limited degree of adaptivity within clearly defined boundaries. Within this framework, the interaction flow is structured as a phase-based progression. The interaction begins with a welcome and orientation phase, designed to promote the child's familiarization with the system and to reduce the emotional impact of entering the interaction. This phase is followed by a decision state that allows the user to choose between two distinct yet complementary pathways: a training-oriented pathway and a play-oriented pathway.

The training phase is intended to gradually introduce the child to the system's functioning, enabling exploration of interaction modalities, comprehension of task requests, and familiarization with dialogic dynamics. This phase also serves an informal diagnostic function, allowing observation of the user's initial level of competence and supporting the subsequent calibration of task complexity. The play-oriented phase, by contrast, constitutes the context in which the educational program proper is initiated, drawing inspiration from the hierarchical structure of the Howlin model for emotion recognition and understanding. In this phase, the child accesses the five levels of the program, each corresponding to a specific level of socio-emotional competence. Each level explicitly defines educational objectives, targeted activity types, permitted conversational strategies, and feedback modalities consistent with the expected developmental stage.

Progression across levels is not automatic but is contingent upon user performance. Advancement to the subsequent level occurs only when the objectives of the current level are achieved; otherwise, the system allows the user to remain within the same state and re-engage with the proposed activities through mechanisms of guided repetition and reinforcement. This iterative logic enables the consolidation of competencies prior to progression, respecting individual learning tempos and supporting a gradual and stable acquisition of emotion recognition skills. Within each state, the conversational agent operates under constrained generative conditions. The language model is employed to formulate questions, prompts, and feedback that are strictly aligned with the pedagogical objectives of the active state, thereby preventing the generation of out-of-context or pedagogically misaligned content. State transitions are governed by explicit rules and may be further supervised by educators, ensuring the maintenance of pedagogical control over the interaction. Moreover, the Howlin program was not directly incorporated into the system but was instead subjected to a process of synthesis and re-elaboration. This process led to the construction of a synthesized representation of the program's core principles, levels, and activities, which served as the conceptual foundation for the design of the state-based interaction tree and for the reconstruction of educational activities in a form compatible with AI-mediated conversational interaction. Finally, the chatbot's behavior is shaped through a retrieval-augmented generation approach, integrated with state-based conversational control and a pedagogically guided structuring of knowledge access.

4. Conclusions

The present contribution proposed a design-oriented framework for the use of Artificial Intelligence as a mediated educational support tool in socio-emotional interventions for children with autism. Within this perspective, the EMORI project integrates consolidated pedagogical principles with artificial intelligence solutions oriented toward interaction control, process predictability, and human supervision. Within this framework, the language model does not operate merely as a response generator, but assumes the role of digital scaffolding, modulating language, task complexity, and feedback strategies according to the active conversational state and the user's responses. A central element of the proposed model is the adoption of a human-in-the-loop approach, which ensures continuous supervision by educators and therapists, preserving pedagogical and therapeutic control over the interaction and preventing unguided automation of educational processes. In line with this approach, the EMORI system is not intended to replace clinical or educational professionals, nor to provide diagnostic or therapeutic decisions, but rather to support structured socio-emotional activities within educational contexts. The present work does not include an empirical validation of the proposed model. Future research will therefore focus on pilot studies and experimental implementations in specialized educational and school settings, with the aim of evaluating system usability, user engagement, and the educational impact of the interactions, as well as examining the role of professional mediation in the integration of Artificial Intelligence into autism-related interventions.

References

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI magazine*, *35*(4), 105-120.

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1, pp. 141-154). Englewood Cliffs, NJ: Prentice hall. Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, 21(1), 37-46. https://doi.org/10.1016/0010-0277(85)90022-8

Cabibihan, J. J., Javed, H., Ang Jr, M., & Aljunied, S. M. (2013). Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *International journal of social robotics*, 5(4), 593-618.. https://doi.org/10.1007/s12369-013-0202-2

Floridi, L. (2022). *Etica dell'intelligenza artificiale: Sviluppi, opportunità, sfide*. Raffaello Cortina Editore.

Goleman, D. (2011). *Intelligenza emotiva. Che cos' è e perché può renderci felici*. Bur.

Holzinger, A. (2016). Machine learning for health informatics. In *Machine Learning for Health Informatics: State-of-the-Art and Future Challenges* (pp. 1-24). Cham: Springer International Publishing.

Howlin, P., Baron-Cohen, S., & Hadwin, J. (1999). *Teoria della mente e autismo-Insegnare a comprendere gli stati psichici dell'altro.* Edizioni Erickson.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in neural information processing systems*, 33, 9459-9474.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA: MIT Press.

Picard, R. W. (2003). Affective computing: challenges. *International journal of human-computer studies*, 59(1-2), 55-64.https://doi.org/10.1016/S1071-5819(03)00052-1

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515– 526. https://doi.org/10.1017/S0140525X00076512

Rudovic, O., Lee, J., Mascarell-Maricic, L., Schuller, B., & Picard, R. W. (2018). Measuring engagement in robot-assisted autism therapy: A cross-cultural study. *Frontiers in Robotics and AI*, 5, 36. https://doi.org/10.3389/frobt.2018.00036

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. https://doi.org/10.1037/h0077714

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9(3), 185–211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Scasellati, B., Admoni, H., & Mataric, M. (2018). Robots for use in autism research. *Annual Review of Biomedical Engineering*, 19, 275–299. https://doi.org/10.1146/annurev-bioeng-071811-150036

Scherer, K. R. (2001). *Appraisal Considered as a Process of Multilevel Sequential Checking*. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 92-120). Oxford: Oxford University Press.

Schiavo, F., Campitiello, L., Todino, M. D., & Di Tore, P. A. (2024). Educational Robots, Emotion Recognition and ASD: New Horizon in Special Education. *Education Sciences*, 14(3), 258. https://doi.org/10.3390/educsci14030258

Vygotsky, L. S. (1978). Mind in society Harvard university press. *Cambridge, MA*.

Wang, K., Sang, G.-Y., Huang, L.-Z., Li, S.-H., & Guo, J.-W. (2023). The Effectiveness of Educational Robots in Improving Learning Outcomes: A Meta-Analysis. *Sustainability,* 15(5), 4637. https://doi.org/10.3390/su15054637

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., ... & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of autism and developmental disorders*, *45*(7), 1951-1966.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, *17*(2), 89-100.