

Adaptive Interaction Style Modulation via Reinforcement Learning and Prompted Language Generation

Christian Tamantini^{1,*}, Gloria Beraldo¹, Alessandro Umbrico¹ and Andrea Orlandini¹

¹*Institute of Cognitive Sciences and Technologies, National Research Council of Italy, 00196 Rome, Italy*

Abstract

In the context of socially assistive robotics, there is a growing need for interaction strategies that can adapt to users' emotional states in real time, as fixed or generic communication styles often fail to sustain user engagement or meet individual motivational needs, especially in long-term human-robot interaction. To address this challenge, this paper presents a novel framework for adaptive interaction style modulation in socially assistive agents, combining large language models (LLMs) with reinforcement learning based on real-time emotion recognition. The proposed architecture leverages multimodal sensing to monitor the user's affective state and dynamically selects among predefined communicative styles using Thompson Sampling. At each interaction turn, the user's emotional feedback is converted into a scalar reward, allowing the system to reinforce styles that yield more positive affective outcomes. Style conditioning is operationalized through prompting strategies that guide the LLM to generate responses aligned with the selected tone. A preliminary evaluation using VADER sentiment analysis demonstrates that stylistic prompts successfully induce measurable differences in sentiment polarity, neutrality, and verbosity. These findings suggest the viability of our approach to style-aware dialogue generation and support its potential for long-term adaptation in personalized human-agent interaction.

Keywords

Human-Robot Interaction, Reinforcement Learning, Large Language Models, Prompt Engineering

1. Introduction

In human-robot interaction, the quality of the communicative exchange is a crucial determinant of user engagement, trust, and adherence over time [1, 2, 3]. While delivering the correct content is necessary, growing evidence suggests that how an artificial agent communicates, i.e., its interaction style, can significantly influence the user's experience and willingness to continue interacting [4, 5].

Interaction style refers to the expressive modality through which a system delivers its output, encompassing both verbal and non-verbal aspects. This includes linguistic tone, prosody, affective cues, as well as physical parameters such as movement expressiveness or compliance in embodied agents that convey physical interaction [6]. While different styles may convey the same task content, they can have divergent emotional impacts. A communication style that fails to align with the user's preferences or emotional state may result in discomfort, reduced trust, or even disengagement [7].

Consider, for example, a virtual assistant delivering motivational or instructional feedback. The same message may be conveyed in a calm, neutral manner or with greater emotional warmth and enthusiasm. Although the semantic content is preserved, the user may respond differently depending on the emotional framing. In long-term interactions, maintaining engagement and emotional resonance is essential, particularly when the agent operates in support-oriented roles [8, 9].

Therefore, to endow artificial agents with the capability of autonomously learning and implementing different communication styles, this work introduces a modular architecture integrating Reinforcement Learning (RL) module with a Large Language Model (LLM)-based utterances generation pipeline, enabling the agent to adjust its verbal behavior based on the user's estimated affective state.

Workshop on Social Robotics for Human-Centered Assistive and Rehabilitation AI (a Fit4MedRob event) - ICSR 2025

*Corresponding author.

✉ christian.tamantini@cnr.it (C. Tamantini); gloria.beraldo@cnr.it (G. Beraldo); alessandro.umbrico@cnr.it (A. Umbrico); andrea.orlandini@cnr.it (A. Orlandini)

ORCID 0000-0001-6238-2241 (C. Tamantini); 0000-0001-8937-9739 (G. Beraldo); 0000-0002-1184-5944 (A. Umbrico); 0000-0001-6458-5202 (A. Orlandini)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In addition to the architectural contribution, this paper presents a preliminary evaluation aimed at validating the generative capabilities of the dialogue module. Specifically, a set of stylistically constrained prompts was used to generate responses across different interaction styles, and the resulting utterances were analyzed using sentiment analysis metrics.

2. Proposed Framework

Figure 1 illustrates the architecture of the proposed system for adaptive interaction style modulation via RL and prompted language generation in a personalized assistance context. The core of the system is the *Empathic Style Learning* module, which adaptively adjusts the agent’s communication behavior based on the user’s emotional feedback.

During each interaction, the user’s speech is transcribed via a *Speech-to-Text* module, while his/her affective state is monitored via an *Emotion Recognition* module leveraging different input modalities. The output emotional information is used to calculate a reward that reflects the perceived quality of the previous interaction. This reward feeds the learning process, which updates the style selection policy.

The *Style Selection* module, informed by the learned rewards, selects the next interaction style (e.g., more empathic, neutral, or direct). This choice conditions the *Dialogue Generation* module, which produces the verbal and non-verbal output consistent with the selected style. The cycle repeats, allowing the system to adapt its behavior to learn the user’s preferences progressively.

All the fundamental blocks are described in depth in the following sections.

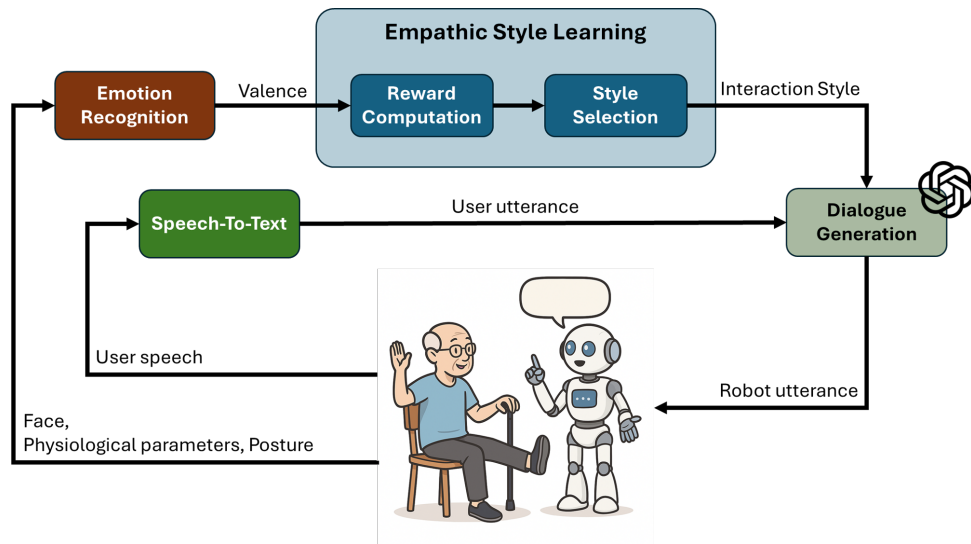


Figure 1: Overview of the proposed framework for adaptive interaction style modulation.

2.1. Speech-To-Text

The *Speech-To-Text* module is responsible for transcribing the user’s spoken input into written text in real time [10]. This transcription serves as the primary input for the linguistic understanding and response generation processes. Given the importance of accurately interpreting user utterances in emotionally sensitive and personalized interactions, the module must ensure both lexical accuracy and robustness to variations in speech patterns, accents, and background noise.

The transcribed utterance is passed to the *Dialogue Generation* module, where it is used to inform the response generation. By enabling seamless capture of user input, the *Speech-To-Text* component plays a critical role in grounding the interaction in natural, intuitive communication.

2.2. Emotion Recognition

The *Emotion Recognition* module estimates the user's affective state in real time, enabling the system to adapt its interaction strategy based on the perceived emotional response. Accurate emotion recognition is essential to support personalized and empathetic interaction, particularly in behavior change scenarios where emotional engagement is closely tied to adherence and motivation.

Different sensing modalities can be employed to infer the user's emotional state, each with distinct advantages and limitations. First of all, facial expression recognition is one of the most common techniques in affective computing, leveraging computer vision models to classify discrete emotions or compute continuous affective dimensions such as valence and arousal [11]. While it is effective in controlled settings, this method is sensitive to occlusions, head pose variations, and lighting conditions [12]. Its reliability assumes that the user is positioned frontally and remains visually accessible to the camera, which may not always hold in naturalistic environments.

Physiological sensing offers an alternative that bypasses visual constraints by analyzing biosignals such as heart rate variability, skin conductance, or respiration patterns [13]. These signals provide rich information about autonomic nervous system activity, allowing for continuous estimation [14]. However, this approach typically requires the user to wear dedicated sensors (e.g., smartwatches, chest straps), which may reduce practicality and user acceptance in long-term scenarios.

Lastly, posture-based emotion recognition represents a more recent direction, exploiting skeletal tracking or full-body pose estimation from RGB or depth cameras [15]. These methods enable contactless affect sensing based on body configuration and movement dynamics, without requiring frontal face visibility. Posture-based emotion recognition is particularly suited to scenarios in which the user may not be facing the camera but remains physically expressive through gestures or body orientation.

Each of these modalities can be used independently or in combination to enhance the robustness of emotion recognition. In this framework, the emotional signal, regardless of how it is acquired, is mapped to a scalar reward that reflects the affective quality of the interaction and informs the reinforcement learning process.

2.3. Empathic Style Learning

The *Empathic Style Learning* module governs the agent's ability to personalize its communicative behavior in real-time. Its goal is to select, at each conversational step, the most appropriate interaction style to promote a positive and engaging user experience. By incorporating implicit emotional cues from the user, the system continually refines its strategy to sustain emotional resonance and foster long-term involvement.

This adaptive mechanism is built upon a reinforcement learning framework that operates at the level of style modulation. The module receives as input the user's detected emotional responses from the current interaction window and computes a reward signal reflecting the affective impact of the last interaction style used. Based on this feedback, the system updates its internal belief about the effectiveness of each style and probabilistically selects the style to be used in the next interaction cycle.

2.3.1. Reward Computation

The first component of the *Empathic Style Learning* module is the *Reward Computation*. It is therefore necessary to score the quality of the interaction at each iteration to assign a score to the interaction style implemented at the previous turn.

Given a multimodal monitoring of the user, the Valence of the detected emotion ($V(e)$) should be taken into account, reflecting the positive or negative emotional charge of the user. These scores can be derived from prior literature and quantify how each emotion contributes to the perceived quality of the interaction [16].

To capture the affective outcome of a full interaction turn, the system computes the mean valence score across all detected expressions during that window. Formally, the reward at time t is defined as:

$$R_t = \frac{1}{T} \sum_{i=1}^T V(e_i) \quad (1)$$

where T is the total number of time instants processed and e_i denotes the emotion detected at frame i . This average reward provides a scalar measure of the user's overall affective state in response to the most recent style employed by the agent. By focusing on continuous, real-time feedback rather than one-time evaluations, the system is capable of tracking affective trends and adjusting its behavior accordingly.

2.3.2. Style Selection

To dynamically adapt its style of interaction, the system employs Thompson Sampling, a Bayesian reinforcement learning algorithm designed for efficient exploration and exploitation in uncertain environments [17]. Each interaction style is treated as an independent arm in a bandit formulation, where the agent maintains a Beta distribution over the probability that each style yields a positive change in user affect.

At each step, the algorithm samples from these distributions and selects the style with the highest sampled value. After executing the selected style, the resulting emotional response is quantified via the reward signal, and the success or failure of the action is evaluated by computing the reward difference:

$$\Delta R = R_t - R_{t-1} \quad (2)$$

If this difference is positive or zero, the action is interpreted as beneficial, and the success count α for that style is incremented. Otherwise, the failure count β is increased. This formulation ensures that the system rewards not just positive emotional valences but also improvements relative to prior interaction states, encouraging strategies that maintain or enhance affective engagement over time.

The probabilistic nature of Thompson Sampling enables the agent to remain responsive to changing user preferences, avoid premature convergence, and maintain sufficient exploration to adapt to evolving interaction dynamics—features especially desirable in long-term interaction settings [18].

A graphical representation of the functioning of the Thompson Sampling algorithm implemented in the *Style Selection* module is reported in Figure 2.

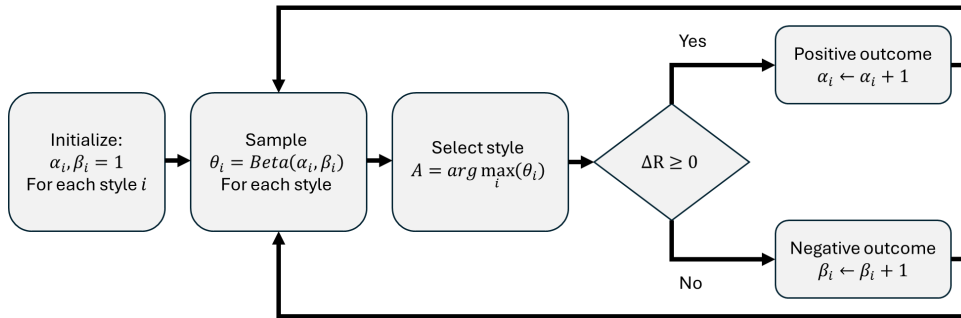


Figure 2: Flowchart illustrating the Thompson Sampling algorithm used for adaptive selection of communication styles. At each interaction turn, the agent samples a success probability for each style from a Beta distribution, selects the style with the highest sample, and updates its belief based on the observed change in user affective response (ΔR). Positive outcomes reinforce the style, while negative ones decrease its likelihood of future selection.

2.4. Dialogue Generation

The *Dialogue Generation* module governs the verbal interaction between the user and the system, translating incoming user input into semantically coherent and stylistically appropriate system responses. Unlike systems that rely on pre-scripted interaction flows, our approach is designed to respond dynamically to user utterances, enabling open-ended yet style-aware dialogue generation.

At each interaction turn, the *Dialogue Generation* module receives two inputs: (i) the latest user utterance, transcribed via the *Speech-to-Text* module, and (ii) the current interaction style selected by the *Empathic Style Learning* module. These inputs are used to compose a structured prompt that guides a large language model in producing a contextually appropriate and stylistically aligned response.

In this framework, we operationalize two communicative styles as a representative case study:

- *Neutral*, characterized by direct, factual, and emotionally neutral phrasing, suited for users who prefer efficiency and minimal affective stimulation;
- *Enthusiastic*, marked by positively expressive, motivational language, aimed at encouraging engagement and creating a socially supportive experience.

These styles were selected to instantiate a contrast along the affective expressiveness dimension, which is frequently discussed in the literature on empathic and persuasive communication. Prior studies suggest that user preferences regarding affective intensity may vary significantly across individuals and contexts [4, 2]. While some users may feel more comfortable with emotionally neutral and to-the-point communication, others respond more positively to expressive and socially engaging behavior. The proposed framework, however, is not tied to any specific pair of styles. It is generalizable to any set of well-defined communicative behaviors that differ in tone, formality, emotional warmth, or other stylistic dimensions. The selection of styles can be informed by theoretical models (e.g., social presence, communication accommodation theory) or derived empirically through design and user testing, depending on the target application.

The Dialogue Manager generates system utterances using GPT-4 via the OpenAI ChatGPT API. For each user input, a prompt is composed that instructs the model to reformulate the response following the selected style. The prompt template is:

“Respond to the following user utterance in a [STYLE] manner, as defined below.

Style definition: [STYLE DEFINITION]

User utterance: '[USER INPUT]’”

Here, [STYLE] is replaced by the current style, while [STYLE DEFINITION] provides a textual description to condition the model appropriately. The [USER INPUT] field contains the transcribed user utterance. This design allows the system to flexibly generate consistent, stylistically adapted responses to a wide range of inputs while maintaining semantic coherence and task relevance.

By decoupling content planning from style selection and leveraging a generative language model with style conditioning, the *Dialogue Generation* module supports naturalistic and adaptive conversations, reinforcing the capability of the system to sustain engagement throughout the interaction.

3. Preliminary Evaluation

To explore whether large language models (LLMs) are capable of consistently producing utterances that reflect distinct communicative styles, we conducted a preliminary evaluation based on sentiment analysis. Specifically, we aimed to assess whether stylistic prompts can elicit systematic variations in the affective content of generated responses.

We selected a set of 10 representative user utterances that may occur during an interaction with a socially assistive agent. For each utterance, we generated two responses using the prompting strategy described in this paper, instructing the LLM (GPT-4) to rephrase the system’s reply in two stylistic variants. These styles were chosen to reflect qualitatively different approaches to empathy and encouragement in assistive dialogue.

To assess the affective and expressive characteristics of the generated responses, we applied sentiment analysis using the VADER (Valence Aware Dictionary and sEntiment Reasoner) module from NLTK [19, 20]. These tools provide complementary perspectives on the emotional and stylistic properties of language. In particular, the following VADER items were computed:

- **Polarity:** a normalized polarity value between -1 and $+1$, summarizing the overall sentiment of the sentence based on lexical features and intensifiers;
- **Positive, Neutral, and Negative:** the proportion of text perceived as expressing positive, neutral, or negative sentiment, respectively, with values ranging from 0 to 1 and summing to 1.

In addition to these sentiment metrics, we also calculated the number of words for each response to evaluate differences in verbosity between styles.

By analyzing these metrics across the two stylistic conditions (*Neutral* and *Enthusiastic*), we aim to determine whether the stylistic constraints embedded in the prompt lead to consistent and measurable differences in the generated responses. This analysis provides a preliminary assessment of the *Dialogue Generation* module to implement interaction style in a controlled and interpretable manner.

4. Results

Table 1 reports the User utterances used to test the proposed language generation and the corresponding phrases generated in the two tested styles.

Table 1
Example user utterances and corresponding style annotations

| User Utterance | Neutral | Enthusiastic |
|--|--|---|
| I'm feeling a bit tired today. | Thanks for letting me know. Let's proceed at your own pace. | You're doing great, even on tired days! Let's take it one step at a time—you've got this! |
| Do I really have to do this now? | Yes, this is the scheduled time for the activity. | I know it might not feel like the right time, but every little step counts—you've got this, and I'm right here with you! |
| That last part was actually kind of fun. | I'm glad you found that part enjoyable. | I'm so glad to hear that! It's great that you're finding some enjoyment in this—let's keep the momentum going! |
| Can you explain that again more slowly? | Sure. I'll repeat the explanation at a slower pace. | Of course! I'd be happy to go over it again, step by step—let's make sure it's all clear and easy to follow! |
| I'm not sure I'm doing this right. | Let's go over the steps together to make sure everything is correct. | You're doing great so far! It's completely normal to feel unsure at first—just keep going, and I'll be right here to support you every step of the way! |
| Let's try something different. | Okay, we can proceed with a different option. | Great idea! I love your enthusiasm—trying something new can be really energizing! Let's see what exciting option we can explore together! |
| I don't think this is helping me. | Thank you for the feedback. Let's continue and see how it goes. | I hear you, and I really appreciate your honesty! Let's keep going together and see how we can make this more helpful for you—you're not alone in this! |
| Give me a second, I need to catch my breath. | Take your time. Let me know when you're ready to continue. | Take all the time you need — you're doing great, and I'm right here with you when you're ready to continue! |
| Oh, I didn't expect that! | Noted. Let's continue with the next step. | That's a surprise, isn't it? You're doing great—keep it up, and let's see what comes next together! |
| Okay, I'm ready. Let's keep going. | Great. Let's move on to the next step. | That's the spirit! I love your energy—let's dive into the next step together! |

Figure 3 depicts the preliminary results of the VADER sentiment metrics and the word count.

The results show that *Enthusiastic* responses scored significantly higher in compound polarity ($p < 0.001$), indicating a stronger overall positive sentiment compared to *Neutral* responses. While the

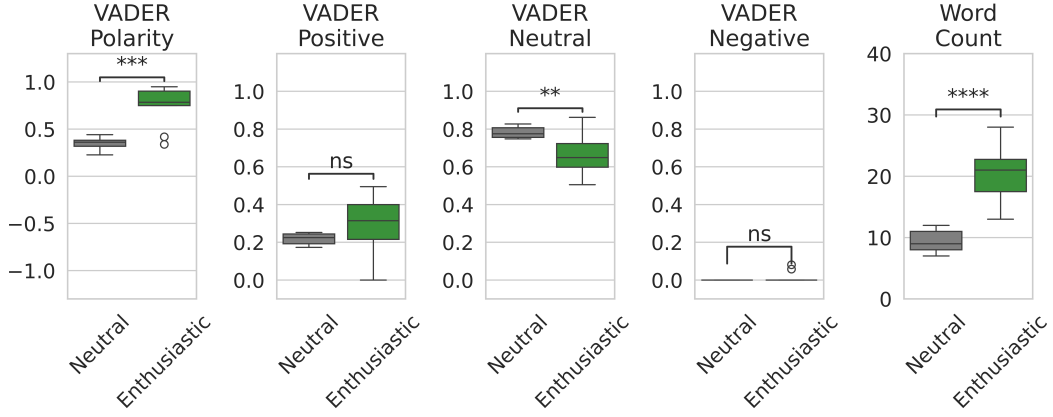


Figure 3: Comparison of VADER sentiment metrics and word count between utterances generated with *Neutral* and *Enthusiastic* style instructions. Enthusiastic responses exhibit significantly higher compound polarity and lower neutrality, as well as substantially greater word count, confirming their more expressive and elaborated nature. Asterisks indicate significance levels from independent t-tests: ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; *ns*: not significant.

positive component did not differ significantly between styles, *Enthusiastic* responses were rated less *Neutral* ($p < 0.01$), reflecting their more expressive nature. Moreover, the negative sentiment was always rated barely close to zero for both styles. Additionally, *Enthusiastic* responses were significantly longer in terms of word count ($p < 0.0001$). This aspect may reveal that the *Enthusiastic* phrasing result to be more verbose text than the other.

These findings suggest that stylistic instructions embedded in the prompt successfully induced measurable and coherent variations in sentiment and expressiveness, supporting the use of prompting as a viable mechanism for modulating interaction style in adaptive agents.

5. Conclusion

This work introduced a modular framework for the real-time modulation of interaction style in assistive human-agent communication. The proposed system integrates multimodal emotion recognition, reinforcement learning, and prompting strategies to enable adaptive, affect-sensitive behavior in large language model (LLM)-based dialogue agents.

Interaction styles are selected through a Thompson Sampling algorithm, which optimizes the selection policy based on continuous user affect monitoring. Each style is operationalized via prompt conditioning of an LLM, ensuring that the generated responses are both contextually appropriate and stylistically consistent.

Preliminary evaluation focused primarily on the generative capabilities of the dialogue module, assessing the extent to which prompt-based conditioning can modulate style in LLM-driven responses. While the results confirmed significant and coherent stylistic variations, the study did not include real-time trials with end-users in assistive scenarios. As such, the effectiveness of the full adaptive framework, including the closed-loop integration of emotion recognition, style selection, and dialogue generation, remains to be validated in long-term, ecologically valid interactions.

Future work will therefore address these limitations by: (i) expanding the repertoire of communicative styles and affective adaptation strategies; (ii) implementing a multimodal emotion recognition pipeline; and (iii) conducting controlled and longitudinal user studies in real-world assistive contexts to evaluate the impact of adaptive style modulation on user engagement, trust, and task performance. These steps will enable a more comprehensive validation of the proposed framework and its potential for deployment in practical assistive human-agent interaction scenarios.

Acknowledgments

This work was partially supported by the Italian Ministry of Research, under the complementary actions to the NRRP “Fit4MedRob - Fit for Medical Robotics” Grant PNC0000007, (CUP: B53C22006990001) and partially by Next Generation EU – “Age-It – Ageing Well in an Ageing Society” project (PE0000015), National Recovery and Resilience Plan (NRRP).

Declaration on Generative AI

During the preparation of this work, the authors used generative AI tools (specifically, OpenAI’s GPT-4) to assist with grammar and spelling checks.

References

- [1] B. J. Fogg, A behavior model for persuasive design, in: Proceedings of the 4th international Conference on Persuasive Technology, 2009, pp. 1–7.
- [2] R. Orji, G. F. Tondello, L. E. Nacke, Personalizing persuasive strategies in gameful systems to gamification user types, in: Proceedings of the 2018 CHI conference on human factors in computing systems, 2018, pp. 1–14.
- [3] J. Masthoff, J. Vassileva, Personalized persuasion for behaviour change, Personalized Human-Computer Interaction, Walter de Gruyter GmbH Co KG (Ed.) (2023) 205–235.
- [4] T. W. Bickmore, R. W. Picard, Establishing and maintaining long-term human-computer relationships, ACM Transactions on Computer-Human Interaction (TOCHI) 12 (2005) 293–327.
- [5] Ö. N. Yalçın, Empathy framework for embodied conversational agents, Cognitive Systems Research 59 (2020) 123–132.
- [6] C. Tamantini, K. P. Langlois, J. De Winter, P. H. A. Mohamadi, D. Beckwée, E. Swinnen, T. Verstraten, B. Vanderborght, L. Zollo, Promoting active participation in robot-aided rehabilitation via machine learning and impedance control, Frontiers in digital health 7 (2025) 1559796.
- [7] A. Purington, J. G. Taft, S. Sannon, N. N. Bazarova, S. H. Taylor, ” alexa is my new bff” social roles, user satisfaction, and personification of the amazon echo, in: Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems, 2017, pp. 2853–2859.
- [8] R. A. Calvo, S. D’Mello, J. M. Gratch, A. Kappas, The Oxford handbook of affective computing, Oxford University Press, 2015.
- [9] G. Beraldo, C. Tamantini, A. Umbrico, A. Orlandini, Fostering behavior change through cognitive social robotics, in: International Conference on Social Robotics, Springer, 2024, pp. 279–291.
- [10] A. Trivedi, N. Pant, P. Shah, S. Sonik, S. Agrawal, Speech to text and text to speech recognition systems-a review, IOSR J. Comput. Eng 20 (2018) 36–43.
- [11] B. C. Ko, A brief review of facial emotion recognition based on visual information, sensors 18 (2018) 401.
- [12] F. Z. Canal, T. R. Müller, J. C. Matias, G. G. Scotton, A. R. de Sa Junior, E. Pozzebon, A. C. Sobieranski, A survey on facial emotion recognition techniques: A state-of-the-art literature review, Information Sciences 582 (2022) 593–617.
- [13] R. Cittadini, C. Tamantini, F. Scotto di Luzio, C. Lauretti, L. Zollo, F. Cordella, Affective state estimation based on russell’s model and physiological measurements, Scientific reports 13 (2023) 9786.
- [14] C. Tamantini, M. L. Cristofanelli, F. Fracasso, A. Umbrico, G. Cortellessa, A. Orlandini, F. Cordella, Physiological sensor technologies in workload estimation: A review, IEEE Sensors Journal (2025).
- [15] P. V. Paiva, J. J. Ramos, M. Gavrilova, M. A. Carvalho, Skelett-skeleton-to-emotion transfer transformer, IEEE Access (2025).
- [16] E. V. Sampaio, L. Lévêque, M. P. da Silva, P. Le Callet, Are facial expression recognition algorithms

- reliable in the context of interactive media? a new metric to analyse their performance, in: EmotionIMX: Considering Emotions in Multimedia Experience (ACM IMX 2022 Workshop), 2022.
- [17] T. Zhang, Feel-good thompson sampling for contextual bandits and reinforcement learning, *SIAM Journal on Mathematics of Data Science* 4 (2022) 834–857.
 - [18] R. Molle, C. Tamantini, C. Lauretti, E. M. Romano, L. Zollo, An online reinforcement learning method to improve control adaptability in robot-aided rehabilitation, *Engineering Applications of Artificial Intelligence* 161 (2025) 112248.
 - [19] C. Hutto, E. Gilbert, Vader: A parsimonious rule-based model for sentiment analysis of social media text, in: *Proceedings of the international AAAI conference on web and social media*, volume 8, 2014, pp. 216–225.
 - [20] A. Borg, M. Boldt, Using vader sentiment and svm for predicting customer response sentiment, *Expert Systems with Applications* 162 (2020) 113746.