

# Development of Adaptable Personality-Orientated Dialogue System for Humanoid Robots

Mayuko Morgan  
Edinburgh Napier University  
mayuko.morgan@napier.ac.uk



## Motivation

In the past five years, there has been a remarkable evolution of humanoid robots[5].

- Historically, humanoid robots are made for specific purposes - e.g. healthcare robots, Pepper.
- Modern trends redefined humanoid robots as general-purpose, versatile, robots capable of engaging with human in socially acceptable way.

However, none possess authentic human-level artificial intelligence or social intelligence akin to that of humans. [1]

Significant disparity between the hype generated by the industry and the widely accepted understanding in research on human-robot interaction.

Adaptability is essential for a robot to be seen as a socially acceptable collaborative team player[2]



Figure 1: Robot receptionists introduced at hospitals in Belgium, by The Guardian, 2016, <https://www.theguardian.com/technology/2016/jun/14/robot-receptionists-hospitals-belgium-pepper-humanoid>

## Public Engagement

How people interact with humanoid robots, offering insight into the different interaction styles exhibited by various age groups and demographics. Some recurring reactions and trends noted, such as frequent instances of initial hesitance towards the appearance of our robots, particularly the very life-like Euclid with some remarks on how eerie and frightening they appear...



Figure 2: Euclid and Moira, developed by Dr. Carl Strathearn

## Personalities and Adaptive Behavior of Robots

### Key for Increasing Users' Acceptance and Trust

Traditional affective robots rely on predefined methods, limiting their adaptability in complex scenarios. LLMs show promise by

enabling context-aware, coherent, and evolving interactions, but their application in robotics is under-explored[8].

## The Big Five - Frequently Adopted Personality-Framework

- The Big Five personality traits are a well-known framework in psychology that categorises human personality into five core dimensions: **Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism**[8].
- LLMs excel at emulating personality based on the Big Five model, but their ability to maintain consistency in multi-turn dialogues decreases with the complexity of the roles[7].
- Robots trained on the big-five may not fully capture the depth and complexity of human emotions, leading to erratic personalities[7].

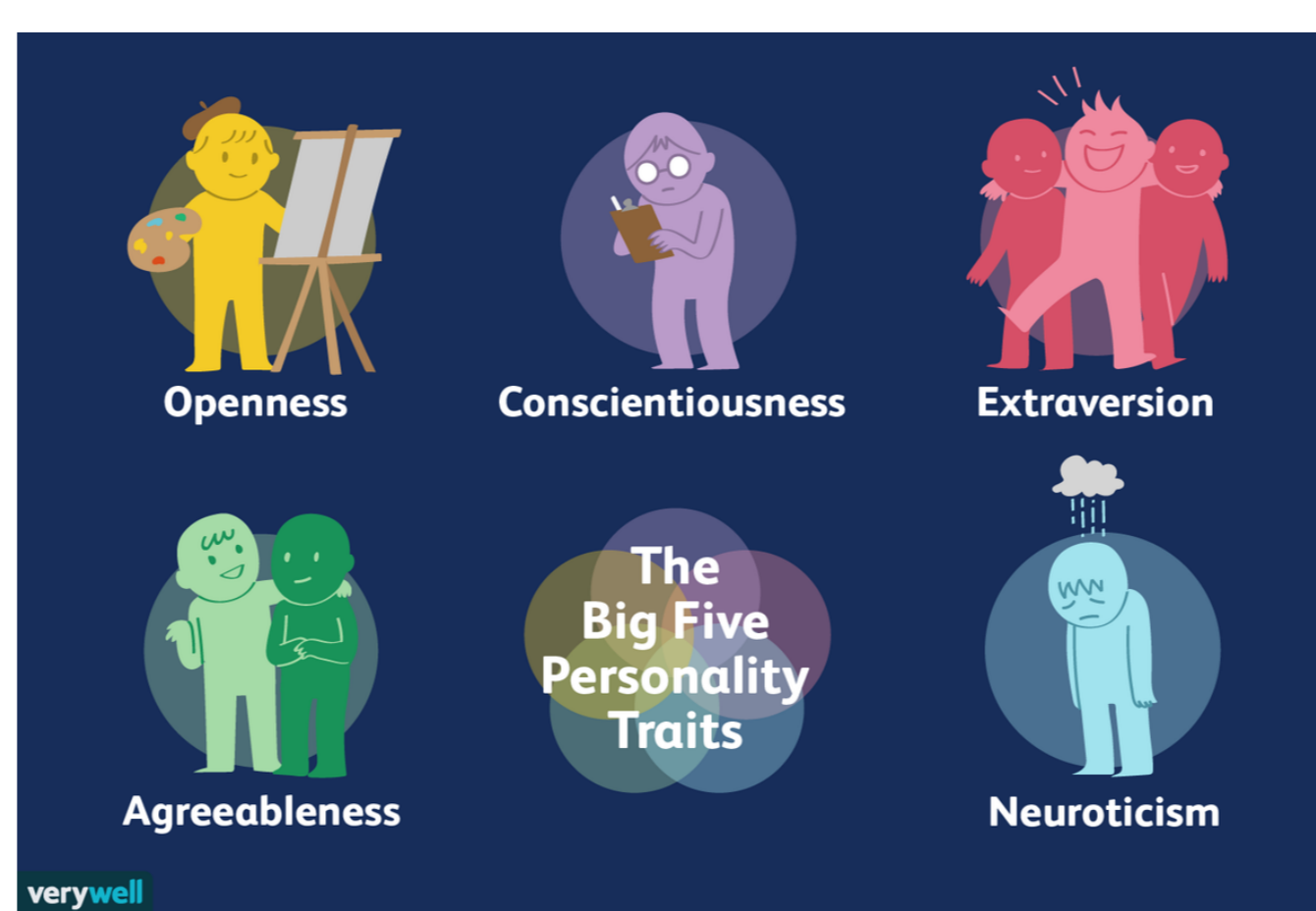


Figure 3: What Are the Big 5 Personality Traits? by verywell mind, 2026, <https://www.verywellmind.com/the-big-five-personality-dimensions-2795422>

## Gartner's Multiple Intelligence Theory

Personality is made up of traits and mental patterns that influence our thoughts and motivations. Howard Gardner's Multiple Intelligence Theory identifies eight distinct types of intelligence: **linguistic, logical-mathematical, musical, bodily-kinesthetic, visual-spatial, interpersonal, intrapersonal and naturalistic**[3].



Figure 4: Gardner's Theory of Multiple Intelligences by verywell mind, 2026, <https://www.verywellmind.com/gardners-theory-of-multiple-intelligences-2795161>

A personality is shaped by our intellectual and cognitive patterns, not just our emotions. I aim to capture the sentiment of each dominant intelligent type across various texts and use these emotional cues to create nuanced personality traits that reflect individual thinking.

## How Can I Capture the Sentiment and Language Style Associated with Each Intelligence Type?

Here are the key three areas of linguistic studies inspired my study:

- Psycholinguistic** - Understand language style varies due to factors like context, author-reader dynamics, and backgrounds[10].
- Computational linguistic** - Promote interdisciplinary research that combines social and computational sciences with the humanities[6].
- Sociolinguistic** - Explore how we use language in different speech community, the situation, the event, the act and the event components, and the functions of speech[4].

The most significant aspect of the initial stage of my research entails thorough selection of transcribed spoken language across different social contexts, circumstances, and speech styles, which will be used to fine-tune the pretrained models.

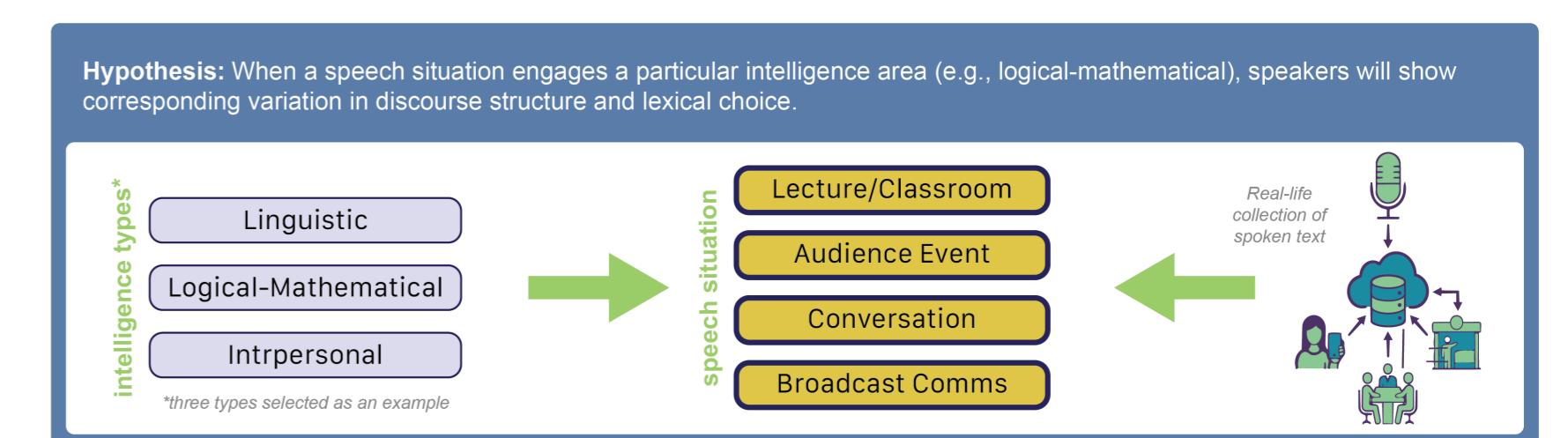


Figure 5: Data Collection

## Controlling the Behaviour of LLM

### Fine-Tuning

Parameter efficient method like LoRa is ideal for modifying specific stylistic elements in text while preserving its core content[9].

### Persona Activation Steering

Neural activations represent the model's evolving "thought process" as information flows through successive layers. Activation steering modifies these representations at inference time to influence behaviour without requiring additional training[10]

## My Approach

Activation steering assumes that a "concept" is represented as a vector in latent space, which can be extracted and injected to influence the LLM's behaviour according to what that vector signifies. Thus, the targeted vector must be "clean" to ensure effective and precise steering[10]. Consequently, my method is to first fine-tune the model to develop a clean, controllable latent representation.

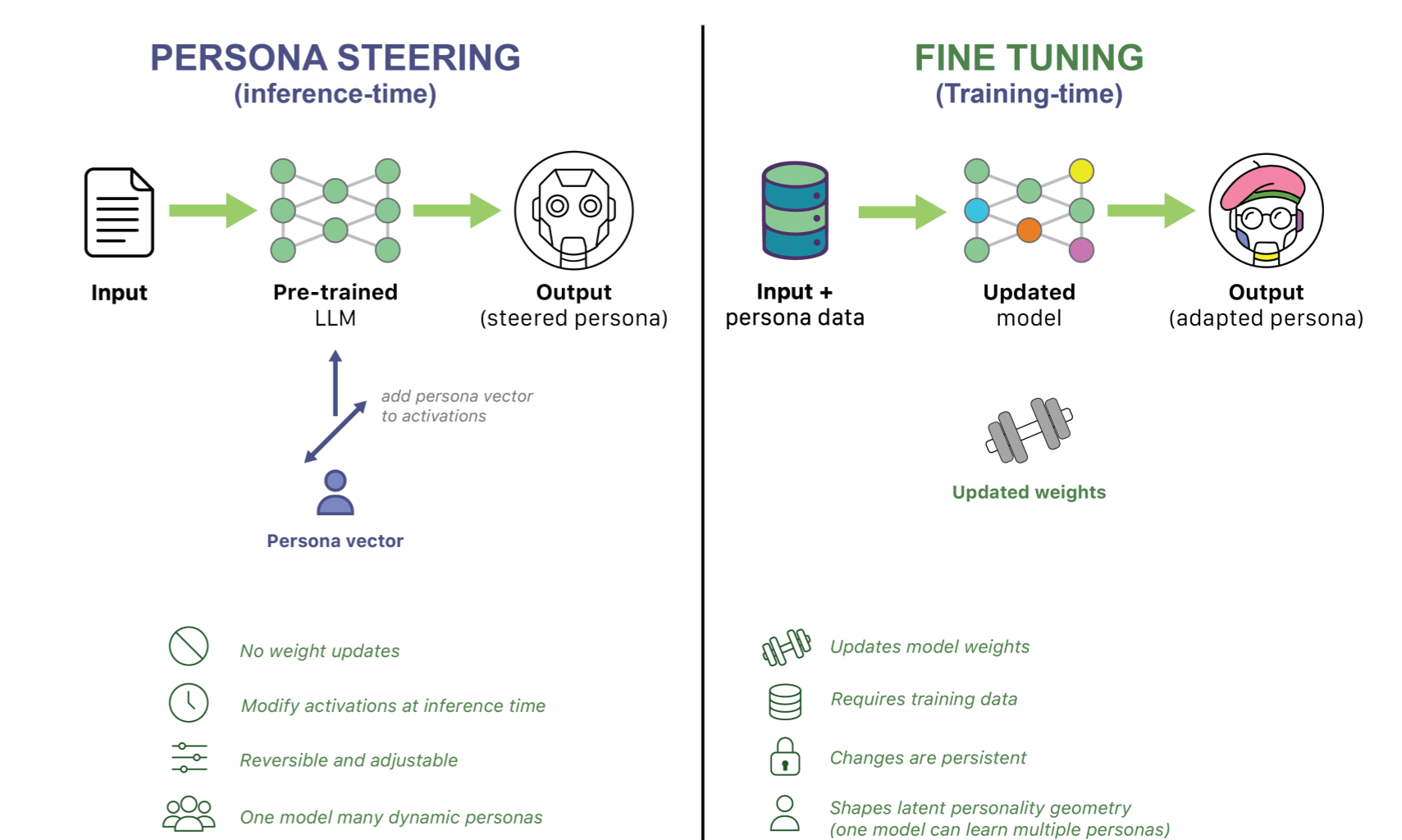


Figure 6: Persona steering and fine-tuning

## References

- Longbing Cao. Humanoid robots and humanoid ai: Review, perspectives and directions. *ACM computing surveys*, 58(4):1–37, 2025.
- Huili Chen, Sharifa Alghowinem, Cynthia Breazeal, and Hae Won Park. Integrating flow theory and adaptive robot roles: A conceptual model of dynamic robot role adaptation for the enhanced flow experience in long-term multi-person human-robot interactions. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI '24*, page 116–126. ACM, March 2024.
- Howard Gardner. *Multiple intelligences: new horizons*. BasicBooks, New York, completely rev. and updat edition, 2006. Includes bibliographical references (p. 263-292) and indexes.
- Dell Hymes. *Models of the Interaction of Language and Social Life: Toward a Descriptive Theory*, chapter 1, pages 4–16. John Wiley Sons, Ltd, 2005.
- Keybotics. History of robotics: A journey of evolution and groundbreaking innovation, 2024.
- Shalom Lappin. What is computational linguistics?, 2014.
- Zhiqiang Ma, Wenchao Jia, Yutong Zhou, Biqi Xu, Zhiqiang Liu, and Zhuoyi Wu. Personality enhanced emotion generation modeling for dialogue systems. *Cognitive computation*, 16(1):293–304, 2024.
- Cheng Tang, Chao Tang, Steven Gong, Thomas M. Kwok, and Yue Hu. Robot character generation and adaptive human-robot interaction with personality shaping, 2025.
- Martina Toshevskaja and Sonja Gievska. Llm-based text style transfer: Have we taken a step forward? *IEEE Access*, 13:44707–44721, 2025.
- Jan Wehner, Sahar Abdelnabi, Daniel Tan, David Krueger, and Mario Fritz. Taxonomy, opportunities, and challenges of representation engineering for large language models, 2025.