

[Research Achievement] Our Faculty Member Awarded Best Paper Award at International Conference!

Unveiling AI's "Emotional Personality": Scientific Comparison of Emotion Understanding Capabilities Across 36 AI Models

We are delighted to announce that Professor Naruki Shirahama from our Faculty of Data Science has been awarded the "Best Paper Award" at The 21st International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2025), held in Taichung, Taiwan from October 15-17, 2025. This prestigious award recognizes groundbreaking research that scientifically elucidated the "emotional individuality" of 36 AI models from seven leading companies through large-scale, standardized evaluation.

The research was presented in the special session "AI-Enhanced QoL, Economics, Affective Computing and Modern Technology," organized by Professor Shirahama himself. The paper was selected as the best paper within the special session from among all accepted papers at the conference, which had an acceptance rate of approximately 60% (86 papers accepted out of 143 submissions).

? Why Was This Research Needed?

Large Language Models (LLMs) such as ChatGPT are rapidly becoming integral to our daily lives, yet the question "Which AI can understand human emotions more accurately?" has remained insufficiently answered.

Previous research primarily focused on evaluating specific AI models individually, making it difficult to fairly compare multiple AIs from different companies using the same criteria. This challenge arose because each company's API specifications differ, making unified evaluation technically challenging.

Furthermore, it remained unclear how emotional evaluations would change when AI models were assigned different "personas" such as "university student," "poet," or "researcher."

Professor Shirahama and his colleagues addressed these challenges by developing a **standardized experimental platform that enables fair comparison of AI models from different**

rent companies on equal footing, conducting the largest-scale comparative study of AI emotion understanding capabilities.



How Was the Research Conducted?

This study conducted a large-scale comparative experiment targeting **36 AI models from seven leading companies** (OpenAI, Google, Anthropic, xAI, DeepSeek, Meta, and Alibaba).

In the experiments, AI models were presented with Japanese literary texts (including "Money and Pistol" by Yumeno Kyusaku and "The Ragged Ostrich" by Takamura Kotaro) and asked to evaluate **how much "Interesting," "Surprise," "Sadness," and "Anger"** they perceived on a 100-point scale.

Furthermore, the study employed a unique methodology by assigning **four different "personas"** to the AI models—"university student," "literary researcher," "poet," and "robot"—and analyzing emotions from each perspective.

The **4,067 data points** collected through this large-scale experiment were analyzed using rigorous statistical methods, including two-way ANOVA and Kruskal-Wallis tests.



Remarkable Discoveries!

This large-scale experiment scientifically demonstrated the following important findings:

1. AI Models Have "Emotional Individuality"!

Statistical verification revealed that **emotional evaluation tendencies differ significantly among AI development companies (vendors)** ($p < 0.05$ across all emotional dimensions).

- **Alibaba models:** Tend to show high evaluations for positive emotions (especially "Interesting")
- **Google models:** Perform balanced evaluations across all emotional dimensions
- **Anthropic models:** Demonstrate highly stable emotional evaluations
- **Meta models:** Display unique evaluation patterns

Even more interestingly, some models exhibited **"fixed individuality,"** maintaining similar evaluations regardless of the assigned persona. Meanwhile, other models acted as

"persona chameleons," flexibly adapting their evaluations according to the given personality.

2. Text Content Strongly Influences Emotional Evaluation!

Exceptionally strong statistical evidence (all $p < 0.001$ with large effect sizes) demonstrated that **textual content accounts for over 50% of variance in emotional evaluations**.

This shows that many AI models can capture and appropriately reflect subtle emotional nuances in literary works. For example, AI evaluations clearly differed between the captivating "Money and Pistol" and the melancholic "The Ragged Ostrich."

3. Highly Reliable AI vs. Still-Developing AI

In this experiment, **models from Anthropic, xAI, and DeepSeek demonstrated perfect reliability, never failing to provide evaluations**. Conversely, some models were unable to return numerical evaluations, revealing areas still under development.

This serves as an important indicator for selecting which models to use when implementing AI in business and services.



Social Significance and Future Prospects

The insights demonstrated through this research provide practical guidelines for model selection when developing **emotion-sensitive AI applications**.

In the future, these findings are expected to lead to the development of services such as:

- **Mental Health Support AI:** Accurately understanding users' emotional states and providing appropriate support
- **Educational Support AI:** Understanding learners' emotional responses and delivering optimal learning experiences
- **Customer Service AI:** Understanding customer emotions and enabling empathetic interactions
- **AI for Literature and Arts:** AI assistants with deeper emotional understanding

Furthermore, students at our university have the opportunity to engage with such **cutting-edge research and develop the skills to create next-generation AI technologies**. Our Faculty of Data Science goes beyond merely acquiring technical skills, actively pursuing research to realize better coexistence between AI and humans.



Comments from Professor Shirahama

"I am deeply honored to receive such a prestigious award. This research represents an important step toward making AI not merely a tool, but a partner capable of understanding human emotions and enabling richer communication.

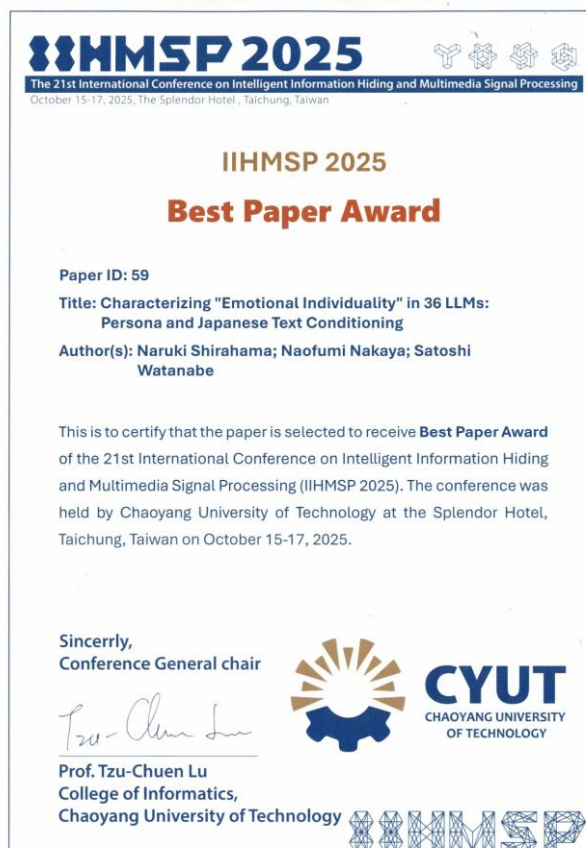
In particular, being able to comparatively evaluate 36 AI models using unified criteria is significant as it overcomes challenges that were difficult in previous research. By scientifically clarifying the 'emotional individuality' that each AI possesses, we can enable optimal AI selection according to specific applications.

I am sincerely grateful to my co-authors, Dr. Nakaya and Professor Watanabe, as well as to the students who continue to learn with great curiosity every day. Data science is a discipline that considers how to utilize the power of AI for society's future. I hope this research achievement resonates with those aspiring to become future data scientists and serves as a catalyst for learning at our university."



Paper Information

- **Paper Title:** Characterizing "Emotional Individuality" in 36 LLMs: Persona and Japanese Text Conditioning
- **Authors:**
 - Naruki Shirahama - Shimonoseki City University, Faculty of Data Science, Professor
 - Naofumi Nakaya - Juntendo University
 - Satoshi Watanabe - Shizuoka Institute of Science and Technology
- **Award:** Best Paper Award
- **Conference:** The 21st International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2025)
- **Venue:** The Splendor Hotel, Taichung, Taiwan
- **Dates:** October 15-17, 2025
- **URL:** <https://iihm25.csie.cyut.edu.tw/>
- **Special Session:** AI-Enhanced QoL, Economics, Affective Computing and Modern Technology
- **Research Support:** This work was partially supported by JSPS Grants-in-Aid for Scientific Research (JSPS KAKENHI), No. 24K13350.



【Contact Information】

Shimonoseki City University
Public Relations and Brand Strategy Division
TEL. +81-83-253-8967