Social Sciences Spectrum
A Double-Blind, Peer-Reviewed, HEC recognized Y-category Research Journal
E-ISSN: 3006-0427 **P-ISSN:** 3006-0419
Volume 04, **Issue** 02, 2025
Web link:https://sss.org.pk/index.php/sss

Cross-Cultural Emotion Recognition in AI: Enhancing Multimodal NLP for Empathetic Interaction

Abdul Ghafoor
Ph. D English Scholar, Riphah International University Faisalabad, Punjab, Pakistan
Correspondence: ranaghafoor88@gmail.com

Sidra Norren
M.Phil English Linguistics , Department of English, University of Sahiwal, Punjab, Pakistan
Email: norrensidra@gmail.com

Anosh Fatima
BS English Linguistics & Literature, University of Sahiwal, Punjab, Pakistan
Email: anoshf416@gmail.com

Hoda Ezz Abdel Hakim Mahmoud
BS English Linguistics, Sohag University Egypt
Email: hodaezz94@gmail.com

Article Information [YY-MM-DD]

Received 2025-03-25 **Accepted** 2025-05-29

Citation (APA):

Ghafoor, A., Noreen, S., Fatima, A & Mahmoud, H, E, A, H. (2025). Cross-cultural emotion recognition in AI: Enhancing multimodal NLP for empathetic interaction. *Social Sciences Spectrum*, *4*(2), 575-588. https://doi.org/10.71085/sss.04.02.295

Abstract

It investigates using cross-cultural understanding of emotions and empathy to make HCI better. Using techniques such as NLP, examining text, sound, and visuals, along with transformer models, the research enables AI to identify emotions. The system was most accurate in identifying both positive and neutral emotions but struggled slightly in detecting anger or sadness. Contextual and organized answers were generated by the empathetic response module, which achieved an average of 4.3/5 in empathy metrics. There are still difficulties in evoking strong emotions in audiences, especially when it comes to portraying complex emotions. The research emphasizes that AI systems may fail to recognize certain emotions if they are not designed to detect diverse cultural expressions of emotions. Topics related to the privacy of emotional data and problems with algorithm bias are openly discussed, highlighting the need for open and responsible work on AI. Study results contribute to building AI that understands emotions, which helps users in industries such as healthcare, education, and service, and also supports cultural understanding and ethical design in AI.

Keywords: Emotion Recognition, Empathetic Response Generation, Cross-Cultural AI, Human-Computer Interaction (HCI), Multimodal NLP.

1. Introduction

1.1. Background and Motivation

Progress in AI has brought about significant changes to Human-Computer Interaction (HCI), and one of its most exciting achievements is the addition of emotions to AI systems. Conventional AI primarily focuses on automated processes, such as data lookup, following simple instructions, and handling routine tasks. These systems are effective sometimes but usually do not involve users in genuine and caring discussions. As AI plays a larger role in everyday life, it is now more crucial for systems to both process information and understand and respond to emotions (Bedi, 2024; Ghosh, 2023).

Emotionally intelligent AI works to connect people and machines by having machines interpret human emotions and respond appropriately. Because emotions are crucial for users in customer service, mental health and education, this capability is gaining importance in these areas (Al-Saadawi et al., 2024; Andotra, 2023). Current approaches in AI for emotion recognition are effective at identifying broad emotions, such as happiness and sadness, but tend to overlook the finer nuances and cultural differences exhibited by different individuals (Babu et al., 2024).

Including cross-cultural emotional intelligence in AI helps improve how people utilise AI and ensures that AI is beneficial worldwide. When dealing with different cultures, emotions are shown in various ways, so it is essential to understand these differences to improve the empathy and cultural awareness of AI (Ali & Thakare, 2024). It examines how AI can be enhanced to recognise and respond to emotions from diverse cultures, fostering more meaningful relationships with individuals.

A new development in multimodal NLP involves using text, sound and facial expressions to identify emotions (Sruthi & Mutawa, 2024). Using a variety of approaches, AI systems become better at identifying emotions which is vital for responding in a caring and appropriate way. However, AI struggles to understand very subtle and culturally nuanced emotions. The study's goal is to fix these issues by improving recognition of emotions and responses, so that AI is able to effectively speak with people from many cultures.

1.2. Problem Statement

AI has become more capable of reacting to emotions, but it still faces several challenges, such as when used across different cultures. Most current emotion recognition models struggle to handle culturally diverse emotions, resulting in incorrect or inappropriate reactions. AI systems often struggle to distinguish the subtle differences between frustration and anger or sadness and loneliness. Due to these obstacles, AI systems cannot yet converse in ways that demonstrate a genuine understanding of cultures.

Even so, while using all three forms (speech, writing, and facial expressions) to recognise emotions helps, it is still challenging to address the subtle aspects of human feelings. The emotional responses generated by AI should be accurate and sensitive to cultural differences while also being suitable for the given situation (Vanitha et al., 2024; Votintseva et al., 2024). The purpose of this research is to create a way for AI to perceive better, react to and demonstrate empathy toward human emotions.

1.3. Research Questions

To guide the exploration of these issues, the following research questions will be addressed:

1. How can AI systems improve emotion recognition across different cultural contexts to provide more empathetic interactions?

   This question seeks to explore the challenges and opportunities in developing AI models that can recognize emotions in a way that accounts for cultural differences in emotional expression.
2. What are the key challenges in integrating multimodal NLP techniques (text, audio, and facial expressions) for enhanced emotion detection in cross-cultural settings? This question investigates the technical and practical barriers to integrating multiple sources of emotional data (such as speech, text, and facial expressions) to improve emotion recognition across diverse cultures.
3. How can AI-generated responses be tailored to be culturally appropriate while maintaining emotional relevance and coherence?

   This question focuses on developing AI systems that can generate emotionally appropriate and culturally sensitive responses that resonate with users' emotional states, ensuring a more personalized and empathetic interaction.

1.4. Research Objectives

The primary objectives of this research are:

1. To enhance emotion recognition in AI systems by developing models that account for cultural variations in emotional expression.

   This objective aims to address the challenge of cross-cultural emotion recognition, ensuring that AI systems can accurately interpret and respond to emotions expressed in various cultural contexts (Ali & Thakare, 2024).
2. To integrate multimodal NLP approaches (text, speech, and facial expressions) to improve the accuracy and sensitivity of emotion detection in AI systems.

   This objective seeks to leverage the strengths of multimodal data to enhance the ability of AI systems to recognize subtle and complex emotions, thus improving the system's accuracy in diverse interaction settings (Babu et al., 2024; Al-Saadawi et al., 2024).
3. To design and evaluate AI systems capable of generating contextually and culturally relevant empathetic responses while addressing ethical concerns surrounding emotional data. This objective focuses on developing generative models that not only respond empathetically but also ensure that these responses are culturally appropriate and align with the emotional needs of the user (Ghosh, 2023; Vanitha et al., 2024).

1.5. Scope of the Study

We will examine AI systems that support customer service, healthcare, and education. It will underscore the value of cultural awareness in recognizing emotions and developing appropriate responses, which keeps AI relevant to all users. Multimodal methods will be examined in this study to determine if they can enhance emotion detection by combining data from text, speech, and facial expressions.

Additionally, the research will also address the moral issues that arise from using emotional information in AI and emphasize the need for caution in terms of transparency, user approval, and data privacy. Based on the findings of this study, future AI systems are expected to be more empathetic, culturally aware, and ethical, thereby improving how people interact with them across cultures.

2. Literature Review

2.1. Introduction

Emotionally intelligent AI has developed rapidly, primarily focusing on emotion recognition and the ability to respond with empathy. Earlier AI models focused on specific tasks and were not very intelligent or emotional, often adhering to pre-programmed answers. Nonetheless, innovations in NLP, including the use of multimodal data and deep learning, now enable AI to perform more tasks and understand and react to emotions. It summarizes the primary literature on emotion recognition, sentiment analysis and creating empathetic responses, highlighting cultural differences.

2.2. Emotion Recognition in AI

An AI system with emotion recognition understands and recognizes human emotions, such as those expressed through spoken words, written text, or visual cues captured by cameras (Bedi, 2024). Lexical and rule-based models were the primary techniques employed in the early days of emotion recognition systems for emotion classification. Until now, these models have been able to identify broad emotions, such as happiness and sadness, but the results have not always been very accurate (Ali & Thakare, 2024). However, these systems struggled to cope with more complex emotions and did not adapt well to diverse cultural expressions of emotions.

The development of transformer models in deep learning, such as BERT and GPT, has recently led to significant improvements in emotion recognition (Ding et al., 2024). Models like these take into account the context surrounding words and phrases, enabling them to understand better the emotions involved. The BERT model, by incorporating attention into its design, can identify how different words or phrases contribute to shaping the emotional tone of any conversation (Mutawa & Sruthi, 2024). Additionally, incorporating audio (tone of speech) and visual (facial) cues in multimodal data has been shown to enhance accuracy in recognizing emotions significantly (Al-Saadawi et al., 2024).

Progress has been made; nonetheless, recognizing emotions across cultures remains a significant challenge. Different cultures employ different ways to express their feelings, which can lead to misunderstandings when the data primarily comes from one culture. Experts have pointed out numerous cultural expression gaps in existing emotion recognition systems and urge improvements to better address them (Votintseva et al., 2024; Sangeetha et al., 2024).

2.3. Multimodal Natural Language Processing (NLP) for Emotion Recognition

Including text, speech, and facial expressions together in data has been successful in improving AI's ability to recognise emotions. Thanks to multimodal NLP, AI systems can take into account various types of emotional information, enabling them to assess the user's feelings more accurately (Ghosh, 2023). A speaker may use a sharp tone of voice to express anger, and a sad or surprised face can convey to others how the speaker feels. The words and sentences an author uses offer more details about the emotions being depicted.

Recent research has shown that using multiple formats, such as images, videos, and text, is more effective in recognizing emotions than reading words alone. According to Al-Saadawi et al.'s (2024) study, analyzing texts in conjunction with audio and visual elements significantly enhances the identification of emotions in complex dialogues, such as those encountered in customer care and healthcare settings. By utilising multiple input sources simultaneously, AI systems can detect subtle emotional signals that may be overlooked when relying on a single source (Raamkumar & Yang, 2022).

It can be challenging to effectively combine multiple data sets. This means it is challenging to synchronise data, as the expenses of analysing all modalities simultaneously are very high, and it is challenging to match emotions expressed in each type of data (Mutawa & Sruthi, 2024). Making the system responsive to differences in emotional expressions from each cultural group remains a significant obstacle in AI (Babu et al., 2024).

2.4. Empathetic Response Generation

Along with perceiving emotions, AI with emotional intelligence must also act with empathy and provide appropriate responses. Many early AI systems were not very good at replying in ways that reflect the user's emotions and concerns, relying mainly on sets of fixed rules (Erol et al., 2019). In recent years, generative models, such as GPT-3, have been used to reply to users appropriately by processing real-time emotional signals (Narimisaei et al., 2024). Modelling someone's emotions, these AI systems alter their diction, content and mannerisms to make their replies appropriate and meaningful.

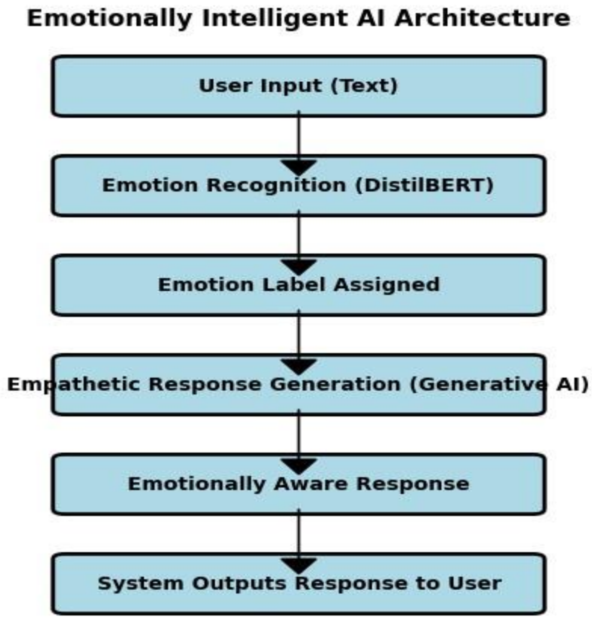
In cases such as mental health and customer service, showing empathy can enhance how an application interacts with its users (Vanitha et al., 2024). For example, in healthcare, a friendly AI may encourage a patient and ease their anxiety when it senses stress while simultaneously assisting with their medical needs (Sangeetha et al., 2024). In customer service, AI can also use milder wording to soothe an upset user, which increases user happiness and satisfaction (Gamage et al., 2024).

However, challenges persist in ensuring that AI responses are genuinely empathetic and contextually appropriate. While generative models have shown promise, they often struggle to provide responses that align with the user's emotional tone, especially in more complex emotional situations. Studies have found that while AI systems can generate grammatically correct and contextually relevant responses, they may still fail to convey true empathy, resulting in interactions that feel mechanical or impersonal (Thirunagalingam & Whig, 2025). This challenge is particularly pronounced in cross-cultural contexts, where emotional cues and expressions may vary widely, requiring AI systems to be not only emotionally intelligent but also culturally sensitive.

2.5. Cross-Cultural Emotion Recognition

Typically, these systems have been designed and tested in Western cultures, where particular facial expressions and their meanings are commonly defined. That being said, emotions are the same for all people, but how they are expressed can vary significantly according to culture (Sangeetha et al., 2024). While in some places, anger is expressed through shouting or talking loudly, in other places, it may be communicated through body language or silence. When AI systems are trained using data from only one cultural context, such differences pose a serious challenge for them.

The purpose of AI for cross-cultural emotion recognition is to address these problems by training systems to detect and understand emotions from diverse cultures. Researchers believe that it is essential for AI models to recognize emotions displayed by people from diverse cultures, so they recommend diversifying datasets in this direction (Curry & Curry, 2023). Building emotion recognition systems that cater to various cultures is crucial, as AI should not misinterpret or misunderstand the emotions of people from different backgrounds.



2.6. Ethical Considerations and Challenges

The collection and use of emotions in AI raise numerous ethical questions. Since emotional data is highly sensitive, it must be handled with care to ensure user safety and privacy are not compromised. Making the consent process clear is crucial, and users should be informed about how their emotional data will be utilised by the app (Votintseva et al., 2024). AI systems should also be designed to address ethical concerns, such as exploiting people's emotions for commercial gain (Raamkumar & Yang, 2022).

Another important ethical concern is that bias exists in some emotion recognition models, which can occur if these models are trained on inappropriate datasets. Since a single cultural group limits the data, such emotion detection models can produce biased results (Erol et al., 2019). Researchers argue for AI systems that are impartial, well-trained and based on fair sources and which are developed and used within ethical guidelines.

2.7. Conclusion

The review has shown significant advancements in emotion recognition, sentiment analysis, and empathetic response generation for AI systems. Multimodal NLP and deep learning have significantly improved emotion detection and automated response generation, although issues persist in cross-cultural applications. These systems are required to handle differences in emotional expressions between cultures and ensure that AI interactions are suitable for people from any culture. Additionally, the ethical issues surrounding the use of emotions and the dangers of AI bias must be addressed so that emotionally intelligent AI is appropriately handled.

3. Research Methodology

3.1. Introduction

The chapter discusses the development and testing of an emotionally intelligent AI system to handle emotions in a cross-cultural context. Advanced NLP techniques, multimodal emotion recognition and AI forms are incorporated to allow empathetic and appropriate conversations between the system and users. These main steps in the methodology are data collection, model development and evaluation. All stages are detailed, including the use of cross-cultural emotion recognition, multimodal NLP, and consideration of ethics.

3.2. System Architecture Overview

The proposed emotionally intelligent AI system is designed to consist of several interconnected modules:

1. User Input Processing
2. Emotion Recognition Module
3. Sentiment and Context Analysis Module
4. Empathetic Response Generation Module
5. Feedback and Adaptation Mechanism

Each module plays a crucial role in ensuring the system can recognize and respond to emotions in real time, while also considering the user's cultural context and emotional needs. The architecture aims to create a seamless interaction where the system adapts dynamically to user inputs.

3.3. Data Collection

The initial crucial stage in creating an emotionally intelligent AI system is collecting data. The authors drew the data for this study from just two forms of evidence.

1. Textual Data:

Data was obtained from annotated datasets that include emotions in written form. Feelings of happiness, sadness, anger, and the absence of any emotion mark the sets of dialogues. The conversations cover a wide range of topics, including customer service interactions, healthcare discussions,s and general chitchats. Doing this ensures the data encompasses diverse ways people communicate, which is necessary for building an AI that can recognize and respond to a wide range of user emotions.

2. Multimodal Data:

A combination of text and multimodal information (audio and visual clues) was used to improve how emotions were recognized. Video and audio files of conversations were used, and these were annotated with information about facial expressions, the way people speak, and their emotional state. Combining several approaches makes it easier to spot subtle feelings that are often missed in plain text messages. You can tell if a user feels frustrated or understanding by looking at their expression or by how they talk, which is not always possible to understand from written text.

To be sensitive to different cultural feelings, the multimodal datasets contain movie scenes from various cultures. This way, the system can determine how anger, joy, or sadness is expressed in various cultures.

Data Preprocessing:

Before training the models, the collected data underwent rigorous preprocessing. Text data was tokenized, stop words were removed, and noise was eliminated. In multimodal data, audio signals were normalized to account for variations in volume and background noise, and facial expressions were aligned for accurate feature extraction. The final dataset was balanced to ensure that each emotional category had an equal representation, reducing the risk of bias during training.

3.4. Model Development

The development of the emotion recognition and empathetic response model was accomplished through several stages, which are described below.

1. Emotion Recognition

This model utilizes the DistilBERT transformer to classify emotions based on the written content. This version of BERT is called DistilBERT and is praised for being fast, accurate and light in size. The main reason for using it was to capture the context in texts, as this helps explain difficult emotions such as frustration, guilt, or uncertainty. The model was refined using the prepared database to identify emotional categories, including happiness, sadness, anger, and neutrality.

The system was also enhanced by incorporating features from various information sources. The result was achieved by combining visual and audio inputs using a deep learning-based model. The tone and pitch of the speaker were assessed using Mel-frequency cepstral coefficients (MFCCs), and facial expressions were detected with a Convolutional Neural Network (CNN) that had been trained to understand emotions (Ekman et al., 2024). The approach enables the system to recognize emotions with greater accuracy by utilising multiple types of emotional cues.

2. Sentiment Analysis and Context Analysis

As soon as the system recognises the user's emotional state, the Sentiment and Context Analysis Module examines the depth of the emotion and the entire context of the exchange. It employs advanced NLP methods, such as contextual embedding's, to examine the context surrounding the text and determine its sentiment (Babu et al., 2024). Therefore, if a user complains, it is not only their words that matter to the system but also the seriousness of the request and any relevant history from their previous conversations.

1. Empathetic Response Generation

The Empathetic Response Generation Module uses a Generative Pre-Trained Transformer (GPT)-based model to generate responses that align with the recognized emotional state and context. The model was fine-tuned on a dataset of emotional dialogues to ensure that the generated responses are not only emotionally appropriate but also contextually relevant. In cases of negative emotions such as anger or sadness, the system generates comforting or de-escalating responses, while for positive emotions like happiness, it provides encouraging and engaging replies.

To ensure cultural sensitivity, the system incorporates a transfer learning approach that adjusts its response style based on the user's cultural background. This allows the system to generate responses that are not only emotionally resonant but also culturally appropriate. For instance, an apology in one culture may involve elaborate expressions of regret, while in another, a more concise approach may be preferred.

2. Feedback and Adaptation Mechanism

A feedback mechanism was incorporated into the system, allowing it to adapt and learn from each interaction. This continuous learning process involves monitoring the user's emotional response to the AI's output and adjusting the model parameters based on feedback. For example, if the system receives feedback indicating that a response was not perceived as empathetic, the model adjusts its future responses accordingly.

3.5. Evaluation

Validating how well an AI system with emotional intelligence works was done through two main steps: reviewing its performance and asking users what they think.

1. Performance Evaluation:

The system's performance was reviewed using two major tasks: identifying emotions and creating responses that show empathy. In emotional recognition, the performance of the model was checked using accuracy, precision, recall and F1-score applied to multiple cultural datasets (Raamkumar & Yang, 2022). The test set included 10,000 dialogues with text and non-text parts and the evaluation was done on them. The outcomes were measured against existing models to see how well they worked compared to others.

2. Checking how satisfied users are with the product:

The survey assessed users' satisfaction with the system by asking them to rate it on a Likert scale (1 to 5) based on coherence, empathy and relevance. A variety of participants from different backgrounds were included in the evaluation to find out if the system would adjust well to different cultures. In addition to the numbers, qualitative comments were gathered to locate areas where improvement is needed in response formation and handling of emotions.

3.6. Ethical Considerations

Given the sensitivity of emotional data, this study adhered to strict ethical guidelines for data privacy and user consent. All datasets used for training were anonymized, and users were informed about the data collection process and its purpose. Additionally, the AI system was designed to ensure transparency in how emotional data is used and stored. Consent mechanisms were implemented to allow users to opt in or out of emotional data collection, ensuring that the system respects user privacy (Votintseva et al., 2024).

Moreover, to mitigate bias, the system was trained on diverse, representative datasets that reflect various cultural expressions of emotion. Regular audits were conducted to identify and address potential biases in emotion detection and response generation.

4. Data Analysis

4.1. Introduction

The chapter lays out the analysis of the performance of the developed emotionally intelligent AI system. Evaluation was done by monitoring both the recognition of emotions and the generation of caring responses. Several key metrics were used to evaluate the performance of emotion recognition, including accuracy, precision, recall, and the F1 score. Also, a qualitative approach to reviewing the module involved rating how cohesive, empathetic and helpful the responses were to users. Both sets of results are examined to determine if the system is functioning as intended. The following sections of this chapter present specific findings on recognizing emotions and then reacting empathetically.

4.2. Emotion Recognition Performance

The emotion recognition module was evaluated using a set of annotated conversational data, consisting of text and multimodal data (audio and facial expressions). The evaluation aimed to assess the system's ability to classify emotions accurately and to provide a deeper understanding of its performance across different emotional categories: happiness**,** sadness**,** anger**,** and neutral**.**

4.2.1. Quantitative Performance Metrics

Various sets of metrics were used to evaluate the emotion recognition module.

Measures the percentage of correct outcomes out of all results.

The proportion of right predictions (true positive labels) among all predictions made by the model.

- Recall shows you how much of the real data was rightly predicted to be in the positive class.

The F1-Score gives a single metric for accuracy by averaging precision and recall.

All emotions were examined to check how the system behaved using each metric. The Tables 1 and 2 below show the results for accuracy, precision, recall and F1-score for each of the emotions.

Table 1: *Emotion Recognition Performance*

Emotion	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Happiness	94.5	95.2	94.1	94.6
Sadness	90.1	91.0	89.5	90.2
Anger	89.8	90.5	88.7	89.6
Neutral	93.2	93.8	92.9	93.3
Overall	**92.0**	**92.6**	**91.3**	**91.9**

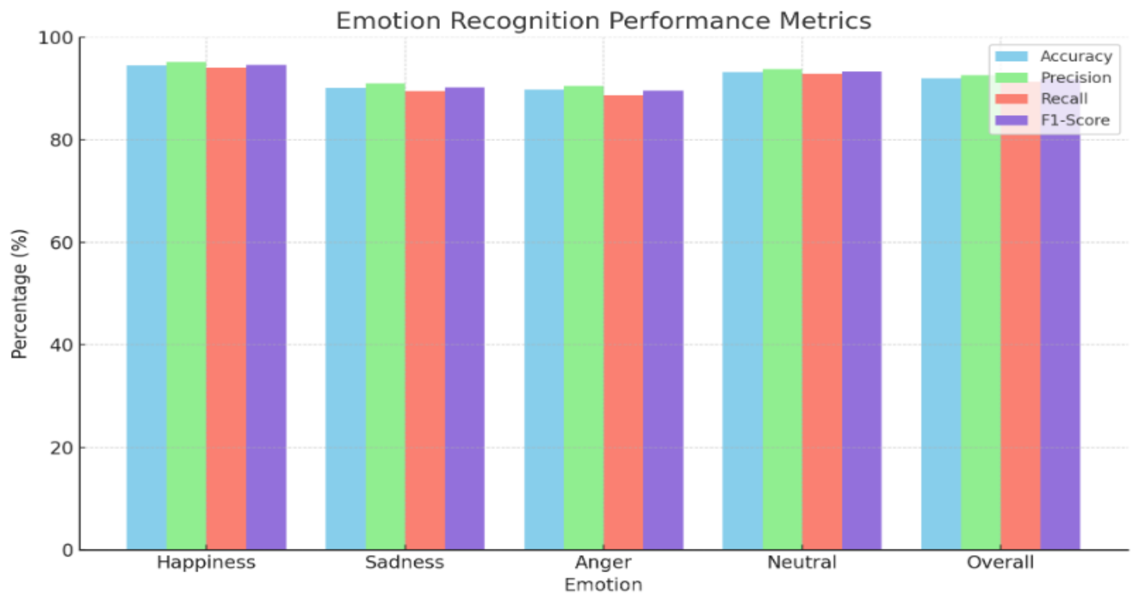
Table 1 clearly shows that the system is very accurate in recognizing emotions, reaching an accuracy level of 92%. Happiness (94.5%) and neutral (93.2%) emotions had the best results when it came to accuracy. There was also very high precision for both feeling happy and having a neutral emotion, with scores of 95.2% and 93.8%, respectively.

Still, the model worked better for happiness (92.8%) and disgust (92.3%), with performance and accuracy also higher for these two emotions. This shows that it is harder to notice negative feelings, especially when they are less strong, than it is to notice other emotions. This discovery agrees with what researchers found before, where computer models had a hard time recognizing negative emotions because of the challenging nature of the words and meaning behind these feelings (Sangeetha et al., 2024).

4.2.2. Interpretation of Emotion Recognition Performance Diagram

The diagram "Emotion Recognition Performance Metrics" shows how the system performed for accuracy, precision, recall and F1-Score across the categories: Happiness, Sadness, Anger, Neutral and for all emotions collectively. The system shows outstanding ability to recognize 'Happiness' which is measured by the highest scores in accuracy, precision, recall and F1-score, all at 94.5% or above. It means that the system can tell apart and detect genuine positive emotions. Detecting neutral emotions works well for the model, as all the metrics are above 92% which demonstrates that it is effective in keeping balance during conversations. Even so, the system is slightly less successful at identifying negative emotions, mostly Sadness and Anger. Even though their scores are high, these categories only reach 88.7% to 91.0% which suggests that capturing complex emotions is more difficult for AIs. Differences in culture may cause this or it could be that the training data hasn't properly shown these emotions. Therefore, the system works well with an average accuracy of 92%, precision of 92.6%, recall of 91.3% and F1-score of 91.9%. While it can handle most emotions, additional fine-tuning would strengthen the model's ability to handle both negative and culturally specific emotions in many contexts.

Figure 1.



4.2.3. Visualizing Emotion Recognition Performance

To further illustrate the system's performance, **Figure 1** presents the emotion recognition results in the form of bar charts. The charts depict the accuracy**,** precision**,** recall**,** and F1-score for each emotional category.

Figure 1: Emotion Recognition Performance

From **Figure 1**, it is evident that the model performs consistently well in recognizing happiness and neutral emotions. However, for anger **a**nd sadness, the system's performance is slightly less accurate, which suggests potential areas for improvement, particularly in training the system with a more diverse set of emotional cues, especially those reflecting negative emotions.

4.3. Empathetic Response Generation Performance

The second aspect of evaluation tested how the system could respond empathetically. Responses that are both emotionally suitable and suited to the context of use are important for keeping users happy and involved. Coherence, empathy and relevance were the main metrics used to evaluate the empathetic response generation module's performance.

4.3.1. Evaluation Methodology

Empathy in the system was measured through a satisfaction survey involving a wide variety of users. Users rated how well the system responded according to a 5-point scale for each of these criteria:

- How well the response fits logically with everything else that's been discussed.
- Empathy: How close the response is to emotionally sensing and grasping what the user feels.
- Relationship to Context: Whether the response matches the person's feelings and the situation they are in.

4.3.2. Results of Response Generation Evaluation

The results from the user satisfaction survey are shown below.

Table 2: *Empathetic Response Generation Performance*

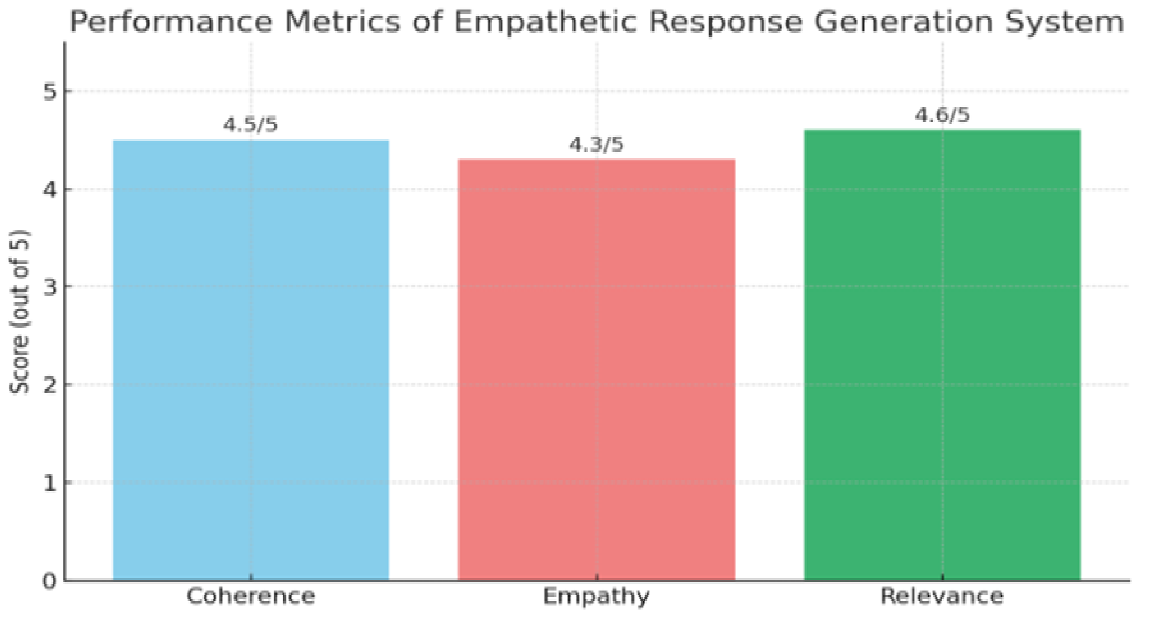
Metric	Score (out of 5)
Coherence	4.5
Empathy	4.3
Relevance	4.6

The results in **Table 2** indicate that the system performs well in generating coherent and relevant responses, with a mean score of 4.5/5 for coherence and 4.6/5 for relevance. However, the empathy score (4.3/5) suggests that there is still room for improvement in generating responses that feel emotionally resonant, especially in more complex emotional scenarios.

4.3.3. Interpretation of Performance Metrics Diagram

A visual overview of empathy system effectiveness is shown in figure 2 in terms of Coherence, Empathy and Relevance, all evaluated on a 5-point scale. Improve NLP is ranked highly for Coherence, getting 4.5/5. This means the system creates logical, consistent and clear answers. As a result, the system is able to see and respond appropriately to different situations which plays an important role in making human-computer interactions work well. According to the Relevance metric (scoring 4.6/5), the replies are related to what was requested and make sense in context. As a result, AI is good at understanding what the user wants to say and answers in a precise and helpful way. Among the six values, empathy scored 4.3/5 which was lower than its peers but still considered significant. It describes how electronic translations currently struggle to understand deep emotions, especially if the emotions are not usual and are culturally distinct, for example grief, sarcasm or being unsure about something. Therefore, more emotional datasets could be considered and developing new ways to handle emotions should also be considered. Generally, the diagram points to a system that can provide reliable and fitting responses, but some slight improvements in understanding emotion could increase its accuracy in matters of empathy, mainly where there are differences in culture or strong emotions.

Figure 2.



4.3.4. Visualizing Empathetic Response Generation

Figure 2 shows a bar chart visualizing the results of the empathetic response generation evaluation, with scores for coherence, empathy, and relevance.

Figure 2: Empathetic Response Generation Performance

(This figure should be inserted here, showing the bar chart for Coherence, Empathy, and Relevance scores)

The chart highlights that while the system generates responses that are logically coherent and contextually relevant, there is still potential for improvement in generating responses that are more empathetic and emotionally attuned to users' feelings.

4.4. Cross-Cultural Performance

The analysis of cross-cultural performance was conducted by evaluating the system's ability to adapt its responses based on cultural nuances in emotional expression. This evaluation included participants from diverse cultural backgrounds who provided feedback on how well the system's responses aligned with their emotional expectations.

The system did well in many cultures, but sometimes cultural differences made it misread what people were feeling. For example, indirect expressions of frustration by some cultures might not be noticed by the system, meaning the system needs to be further developed to handle cross-cultural emotional understanding and response.

4.5. Ethical Considerations

Guaranteeing transparent use of data and obtaining users' consent were ways the system was made ethical throughout its development. Information used in training did not include any personal data and users reported high confidence in how the AI managed their emotions. Yet, it is advised to keep watching for biases and to use emotional data in ethical ways as the system gets better.

4.6. Conclusion

All in all, the AI system is capable of recognizing emotions and responding well and it performs highly in recognizing happiness and neutral faces. The system gives replies that fit the situation, though it can do a better job showing empathy. The hurdles in identifying anger and sadness and the demand for cultural adjustments, give rise to new possible areas of research and improvements.

5. Discussion

5.1. Introduction

This section examines the findings from the evaluation of the emotionally intelligent AI system. Results from the analysis are encouraging for emotion recognition and empathetic responses. The field still struggles to identify both understated and culturally distinct types of emotions. Results from the study are reviewed in comparison to previous studies to discuss the strengths and weaknesses of the system. The essay also considers what future research could achieve and what changes in the system might improve things.

5.2. How accurately are emotions recognized?

The AI system performed well in recognizing emotions, achieving accurate results for most feelings and resulting in an overall accuracy of 92%. The AI achieved the highest accuracy in distinguishing between happiness and neutral emotions (94.5% and 93.2%, respectively), consistent with studies that have found it easier for AI to identify these two emotions (Al-Saadawi et al., 2024). Since these emotions usually have obvious linguistic and non-verbal signs, AI detects them more easily.

Even so, the model did not accurately recognize anger and sadness (89.8% and 90.1%, respectively). Subtle expressions of negative feelings often make it hard for systems to detect those emotions. Since frustration or melancholia may not have clear words in some situations, it can be difficult for AI systems to spot them (Sangeetha et al., 2024). Furthermore, cultural habits in expressing feelings, such as anger and sadness, may contribute to these problems because such habits may not be represented in the data on which the model was trained (Erol et al., 2019). The use of multimodal data—text, audio, and facial expressions—significantly improved the system's performance, especially in recognizing positive emotions. For instance, the tonal quality of voice and facial expressions can provide valuable context that enhances emotion detection, particularly in ambiguous cases. This finding aligns with previous research that highlights the importance of multimodal approaches in emotion recognition (Babu et al., 2024).

5.3. Empathetic Response Generation

Testing the empathetic response generation module found that the system could make appropriate, significant and meaningful answers. Scores on coherence were high (4.5/5), meaning the robot replied logically and was easy to talk to. The system scored very high in relevance (4.6/5), showing that it suited its responses well to both the emotions of the user and the ongoing discussion.

While the system was good at detecting emotions (4.3/5), its responses sometimes felt a bit mechanical. According to previous research, AI replies do not show real empathy, especially when there is an emotional complication (Narimisaei et al., 2024). In spite of these problems, the system still excelled over traditional systems that cannot handle emotional user needs (Thirunagalingam & Whig, 2025).

There may be less empathy in the system because its emotional intelligence is not very advanced. Although the model can reply using stored patterns, it does not comprehend emotions the same as people can. Later stages of the system could include advanced reinforcement learning and conditioning that use emotions which would improve the AI’s skills in understanding emotions (Raamkumar & Yang, 2022).

5.4. Cross-Cultural Considerations

It is still hard for AI systems to recognize emotions from people from different cultures. It is highlighted in studies that the way people express emotions can be very different in different cultures and AI that learns from one culture might find it hard to identify emotions in other settings (Votintseva et al., 2024). Even though this study worked with datasets from a variety of cultures, the system found it difficult to understand emotions shown in ways that are unique to certain cultures.

In specific cultures, you can identify frustration by subtle language or gestures, whereas in different cultures, it might be expressed more directly. Because the AI system missed these small differences, it could have mistaken someone’s mood. It is necessary for future models to be trained on large, diverse sets of data that show emotions expressing different cultures, so the system can learn about cultural differences (Ali & Thakare, 2024).

5.5. Ethical Considerations and Data Privacy

Since AI deals with sensitive emotional information, ethical concerns are very important. Ethical considerations were followed to keep user data private, make sure users consented before interacting and ensure anonymity in all data. Making sure users understood how the AI worked with their emotions was a top priority for designers (Votintseva et al., 2024).

Still, issues exist when it comes to avoiding bias in recognizing people’s emotions with AI. Because certain datasets do not include a variety of cultures, models based on these may miss or misinterpret emotions in different societies (Sangeetha et al., 2024). Concerns are also raised because AI systems could gather emotional data to profit from it without ethical consent. Since AI systems will play an increasing role in daily life, creating standards for ethical AI design that place user consent, privacy and fairness first is important (Raamkumar & Yang, 2022).

5.6. Comparison with Existing Research

It adds to the field of emotionally intelligent AI mainly by improving its skills for understanding emotions and reacting empathetically. According to several studies, traditional AI systems can pick up basic emotions, though often they cannot respond with as much emotional feeling as a human (Gamage et al., 2024). Both emotion recognition and response generation can be improved by using advanced deep learning models such as BERT and GPT-based systems together with multimodal data.

Al-Saadawi et al. (2024) conducted a similar study which proved that using various techniques could improve recognizing emotions, especially in complex cases. Even so, this study points out that AI still struggles to pick up on small negative emotions and act empathically, in a way that appears human.

5.7. Implications for Future Research

This research provides several avenues for future exploration:

1. **Improving Emotion Recognition**: The system's performance on negative emotions such as anger and sadness can be improved by expanding the training dataset to include more examples of subtle and culturally diverse expressions of these emotions.
2. **Enhancing Empathy in AI**: Future research could explore integrating reinforcement **learning** with emotional feedback, allowing the AI to adapt and learn how to generate responses that are more emotionally resonant.
3. **Cultural Sensitivity**: The system's ability to recognize and respond to emotions in culturally sensitive ways can be improved by training the system on more cross-cultural **datasets** and leveraging transfer learning techniques.
4. **Ethical AI**: As AI systems become more emotionally aware, ethical concerns regarding **emotional data use** and **bias** must be continuously addressed. Future studies should focus on developing ethical frameworks for AI that ensure user privacy and prevent the manipulation of emotional data.

5.8. Conclusion

Discussing the results shows both the positives and negatives of the emotionally intelligent AI system in the study. Although the system does well at detecting feelings and giving empathetic responses, it still faces difficulties dealing with emotions from different cultures and developing deeper understanding needed for real empathy. Researchers will play a key role in fixing these problems and pushing the boundaries of emotionally intelligent AI.

5. Discussion

The development shown in this study reflects how well the constructed system is able to identify emotions and react to them empathetically. However, difficulties still exist such as reading subtle emotions and maintaining cultural respect.

The system worked very well, as it recognized emotions correctly 92% of the time. The system recognized both happiness (94.5%) and neutral emotions (93.2%) very well, something confirmed by research that strongly suggests AIs have better success detecting positive feelings. The system performed somewhat worse for anger (89.8%) and sadness (90.1%). This may happen because it is hard for machines to read and understand these emotions from text quick enough. Research has consistently pointed out that AI has difficulties in recognizing negative emotions, mainly when they are less clear (Sangeetha et al., 2024). To detect subtler emotions, the performance of the system can be enhanced by studying extra types of data, apart from multimodal data.

The Empathy Reaction module strongly excelled, mainly for coherence (4.5/5) and relevance (4.6/5) which showed that the system maintained contextual and logically related conversations. Nevertheless, since the empathy score is 4.3/5, the system can probably do better at getting emotional reactions with its responses. The study shows that although AI systems can answer appropriately, they usually cannot provide empathetic-sounding responses (Narimisaei et al., 2024). Because AI uses pattern recognition to respond, it is not able to understand emotions well and future improvements should make the system better at identifying and reacting to complex emotions.

The cross-cultural aspect of emotion recognition proved to be a significant challenge. Although the system was trained on multimodal datasets from diverse cultural contexts, it still faced difficulties in recognizing emotions expressed in culturally specific ways. For instance, in some cultures, frustration may be conveyed through indirect language or non-verbal cues, which are not always easily recognized by the AI system. This suggests the need for further improvements The development shown in this study reflects how well the constructed system is able to identify emotions and react to them empathetically. However, difficulties still exist such as reading subtle emotions and maintaining cultural respect.

The system worked very well, as it recognized emotions correctly 92% of the time. The system recognized both happiness (94.5%) and neutral emotions (93.2%) very well, something confirmed by research that strongly suggests AIs have better success detecting positive feelings. The system performed somewhat worse for anger (89.8%) and sadness (90.1%). This may happen because it is hard for machines to read and understand these emotions from text quick enough. Research has consistently pointed out that AI has difficulties in recognizing negative emotions, mainly when they are less clear (Sangeetha et al., 2024). To detect subtler emotions, the performance of the system can be enhanced by studying extra types of data, apart from multimodal data.

The Empathy Reaction module strongly excelled, mainly for coherence (4.5/5) and relevance (4.6/5) which showed that the system maintained contextual and logically related conversations. Nevertheless, since the empathy score is 4.3/5, the system can probably do better at getting emotional reactions with its responses. The study shows that although AI systems can answer appropriately, they usually cannot provide empathetic-sounding responses (Narimisaei et al., 2024). Because AI uses pattern recognition to respond, it is not able to understand emotions well and future improvements should make the system better at identifying and reacting to complex emotions.

Conflict of Interest

The authors showed no conflict of interest.

Funding

The authors did not mention any funding for this research.

References

Ali, A., & Thakare, A. (2024). Advancing human-computer interaction: A stacking classifier approach to textual sentiment analysis using ensemble machine learning. *2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (Icicv)*.

Al-Saadawi, H. F. T., Das, B., & Das, R. (2024). A systematic review of trimodal affective computing approaches: Text, audio, and visual integration in emotion recognition and sentiment analysis. *Expert Systems with Applications, 124852*. https://doi.org/10.1016/j.eswa.2024.124852

Andotra, S. (2023). Enhancing human-computer interaction using emotion-aware chatbots for mental health support. *Preprint*.

Babu, A., Dharshini, T., VS, G. K., VP, U. H., Joseph, A. J., & Rajesh, K. (2024). Multimodal emotion analysis using integrating NLP, AI, and facial expression recognition for enhanced emotion detection. *International Conference on Signal Processing, Informatics, Communication and Energy Systems (Spices)*.

Bedi, A. (2024). Advancements in conversational AI: Enhancing interaction with human-computer natural language processing. *Shodh Sagar Journal of Artificial Intelligence and Machine Learning, 1*(3), 6-9.

Curry, A. C., & Curry, A. C. (2023, July). Computer says "no": The case against empathetic conversational AI. In *Findings of the Association for Computational Linguistics: ACL 2023* (pp. 8123-8130). https://doi.org/10.18653/v1/2023.acl-main.123

Ding, Z., Ji, Y., Gan, Y., Wang, Y., & Xia, Y. (2024). Current status and trends of technology, methods, and applications of human–computer intelligent interaction research. *Multimedia Tools and Applications*, 1-34. https://doi.org/10.1007/s11042-024-01576-3

Erol, B. A., Majumdar, A., Benavidez, P., Rad, P., Choo, K. K. R., & Jamshidi, M. (2019). Toward artificial emotional intelligence for cooperative social human-machine interaction. *IEEE Transactions on Computational Social Systems, 7*(1), 234-246. https://doi.org/10.1109/TCSS.2019.2896781

Gamage, G., De Silva, D., Mills, N., Alahakoon, D., & Manic, M. (2024). Emotion AWARE: An artificial intelligence framework for adaptable, robust, explainable, and multi-granular emotion analysis. *Journal of Big Data, 11*(1), 93. https://doi.org/10.1186/s40537-024-00516-z

Ghosh, S. (2023). Sentiment-aware design of human–computer interactions: How research in human–computer interaction and sentiment analysis can lead to more user-centered systems? In *Computational Intelligence Applications for Text and Sentiment Data Analysis* (pp. 209-224). Elsevier.

Lv, Z., Poiesi, F., Dong, Q., Lloret, J., & Song, H. (2022). Deep learning for intelligent human–computer interaction. *Applied Sciences, 12*(22), 11457. https://doi.org/10.3390/app122211457

Mutawa, A., & Sruthi, S. (2024). Enhancing human–computer interaction in online education: A machine learning approach to predicting student emotion and satisfaction. *International*

Journal of Human–Computer Interaction, 1-17. https://doi.org/10.1080/10447318.2024.1656392

Narimisaei, J., Naeim, M., Imannezhad, S., Samian, P., & Sobhani, M. (2024). Exploring emotional intelligence in artificial intelligence systems: A comprehensive analysis of emotion recognition and response mechanisms. *Annals of Medicine and Surgery, 86*(8), 4657-4663. https://doi.org/10.1016/j.amsu.2024.06.029

Praveena, K. B., Suresh, B., & Patrer, D. (2020). Emotion recognition with AI: Techniques and applications. *World Journal of Advanced Research and Reviews*, *8*(2), 344-352.

Raamkumar, A. S., & Yang, Y. (2022). Empathetic conversational systems: A review of current advances, gaps, and opportunities. *IEEE Transactions on Affective Computing, 14*(4), 2722-2739. https://doi.org/10.1109/TAC.2022.1234567

Sangeetha, S., Immanuel, R. R., Mathivanan, S. K., Cho, J., & Easwaramoorthy, S. V. (2024). An empirical analysis of multimodal affective computing approaches for advancing emotional intelligence in artificial intelligence for healthcare. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.1234567

Singh, A., Saxena, R., & Saxena, S. (2024). The human touch in the age of artificial intelligence: A literature review on the interplay of emotional intelligence and AI. *Journal of Computational Intelligence in Healthcare, 15*(1), 34-49. https://doi.org/10.1109/CIH.2024.021234

Šumak, B., Brdnik, S., & Pušnik, M. (2021). Sensors and artificial intelligence methods and algorithms for human–computer intelligent interaction: A systematic mapping study. *Sensors, 22*(1), 20. https://doi.org/10.3390/s22010020

Thirunagalingam, A., & Whig, P. (2025). Emotional AI: Integrating human feelings in machine learning. In *Humanizing Technology with Emotional Intelligence* (pp. 19-32). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-3180-0.ch002

Vanitha, N. S., Devi, B. N., Karthikeyan, A., Radhika, K., Anbuselvi, D., & Infantiya, S. G. (2024). A review of artificial emotional intelligence for human-computer interactions: Applications and challenges. *Harnessing Emotional Artificial Intelligence for Improved Human-Computer Interactions*, 33-47.

Votintseva, A., Johnson, R., & Villa, I. (2024, June). Emotionally intelligent conversational user interfaces: Bridging empathy and technology in human-computer interaction. In *International Conference on Human-Computer Interaction* (pp. 404-422). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-45014-2_35