Enhancing User Experience in Chinese Initial Text Conversations with Personalised AI-Powered Assistant

Jindi Wang
jindi.wang@durham.ac.uk
Durham University
UK

Ioannis Ivrissimtzis
ioannis.ivrissimtzis@durham.ac.uk
Durham University
UK

Zhaoxing Li
zhaoxing.li@soton.ac.uk
University of Southampton
UK

Lei Shi
lei.shi@newcastle.ac.uk
Newcastle University
UK

ABSTRACT

ABSTRACT

In the rapidly evolving landscape of text-based communication, the importance of the initial interaction phase remains paramount. This study investigates the potential benefits that a proposed AI chat assistant equipped with text recommendation and polishing functionalities can bring during initial textual interactions. The system allows the users to personalise the language style, choosing between *humorous* and *respectful*. They can also choose between three different levels of AI extraversion to suit their preferences. Results of user evaluations indicate the system received a “good” usability rating, affirming its effectiveness. Users reported heightened comfort levels and increased willingness to continue interactions when using the AI chat assistant. The analysis of the results offers insights into harnessing AI to amplify user engagement, especially in the critical initial stage of textual interaction.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Empirical studies in HCI**.

KEYWORDS

Chat assistant, Personalisation, Conversational AI, Computer-Assisted Human Interaction

ACM Reference Format:
Jindi Wang, Ioannis Ivrissimtzis, Zhaoxing Li, and Lei Shi. 2024. Enhancing User Experience in Chinese Initial Text Conversations with Personalised AI-Powered Assistant. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA ’24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3613905.3651104

1 INTRODUCTION

Large language models (LLMs) like GPT-4 [30], have demonstrated impressive capabilities in comprehending and generating text in ways that mirror closely human communication. These AI models can be conditioned to mimic specific writing styles, thereby enabling possibilities of style-specific text generation for refined digital communications [20]. Moreover, AI’s ability to discern sentiment from text [36] can be utilised to understand the emotional undertones in online conversations and respond with empathy, a critical aspect towards an enriched user experience.

AI-powered chat assistants have emerged as a transformative force in the field of computer-assisted human interaction [39]. This revolution is not merely about convenience or efficiency; it touches upon the foundational aspects of human interaction, such as comfort and willingness to engage. That is, AI-powered conversational interfaces hold the promise, not only of enhanced utility but also to facilitate deeper human connection [32].

One critical stage in both digital and face-to-face conversations is the initial interaction, often termed “icebreaking”. Despite being a brief phase, it can be delicate and set the tone for the entire communication, affecting the trajectory of subsequent engagement [33], in both professional and informal environments. Thus, given the persistent availability of AI chat assistants and their ability to process vast amounts of information swiftly [10], integrating them into these initial phases can have profound implications for user experience. In particular, functionality such as text *recommendation* and *polishing* can refine and guide these interactions, making them more meaningful and user-centric [19].

While a great deal of research on Conversational AI interfaces has focused on their applications in task-oriented scenarios, such as customer support and virtual assistance [5], their potential to enhance human-human conversations has been relatively underexplored. This scarcity of relevant research is even more pronounced on the more specific topic of AI-assisted initial human-human interactions.

Thus, recognising the pivotal role of initial text conversation in shaping subsequent interactions, and noting that no currently-used chat application incorporates a personalised AI assistant, this study focuses on this specific application domain. It aims to address the primary research question: ***In initial textual interaction scenario, can an AI chat assistant, equipped with text recommendation and polishing functionalities, increase the user’s comfort level and willingness to continue the conversation?*** We believe that understanding the influence of AI in initial user interactions is crucial for designing user-centric, AI-driven chat interfaces.

CHI EA ’24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05.
https://doi.org/10.1145/3613905.3651104

2 RELATED WORK

2.1 "Breaking the Ice" in Digital Communication

The act of "breaking the ice" in digital communication has been a topic of interest in the field of Computer-Human Interaction (CHI) and related disciplines. The transition from face-to-face to digital communication has brought unique challenges and opportunities in establishing initial rapport. Hancock *et al.* [15] noted that the absence of non-verbal cues in text-based communication can both hinder and facilitate self-disclosure, depending on the context. Tidwell and Walther [35] found that, in online settings, users often employ strategic self-presentation techniques to "break the ice" effectively. These strategies might involve sharing more personal information earlier in the conversation than in face-to-face interactions. Another variable is the role of the interface design in easing initial interactions. Vlahovic *et al.* [37] highlighted that platforms incorporating gamified elements or icebreaker questions can significantly reduce the initial tension and promote organic conversations. As digital communication platforms evolve, understanding and enhancing the initial text conversation phase remains pivotal for fostering meaningful online connections.

2.2 Text-based Communication Language Style

The language style in text-based communication has been thoroughly studied from various angles, including sociolinguistics, computational linguistics, and computer science.

2.2.1 ***Sociolinguistic Perspectives.*** Text-based communication merges speech and writing elements, creating a unique form that combines the informality of spoken language with the permanence of written text [3, 17]. The format demands that users possess "communicative competence" [21], which is the ability to communicate effectively within specific contexts. Core to its style are stylistic variations, such as abbreviations, emojis, and non-standard punctuation [7, 34], which often convey aspects of the user's identity like age and dialect [18]. Here, however, our chat assistant maintains a standardised format by not including these stylistic elements. Emotional expressions, ranging from capitalisations for emphasis to emojis for nuanced feelings, enrich this communication form [22]. Herring *et al.*'s categorisation for Computer-Mediated Discourse [16]—comprising system, participation, structure, and tone dimensions—aptly captures text communication's essence. The tone, varying from formal to playful, significantly influences the user experience and communication efficiency.

2.2.2 ***Computer Science Perspectives.*** Text-based communication on digital platforms, characterised by its distinct linguistic attributes such as brevity and informality, poses both challenges and opportunities for the fields of NLP, HCI, and dialogue systems. Traditional NLP techniques often struggle with non-standard language forms [8], prompting researchers to adapt techniques specific to platforms, as exemplified by the need for tailored POS tags in Twitter [13].

Out-of-vocabulary (OOV) words are prevalent in social media text, and they pose significant challenges [14]. Furthermore, the evolving nature of online language necessitates periodic model updates [6]. Emoticons and unconventional punctuation, while complex, offer opportunities for nuanced language models to excel [27].

Demographic factors, including age [18], gender [1], and location [9], have been identified as influential factors in online linguistic styles. Recent advancements in the domain focus on style adaptation in text generation, with methods ranging from unsupervised techniques [24] to back-translation approaches [29]. The intersection of personalisation and style is emerging as a critical area of study, highlighting the significance of individual stylistic preferences in enhancing user-system interactions [25, 28].

2.3 Personalisation and Style Adjustment in Chat Assistants

2.3.1 ***Personalised Response Generation.*** Recent research has adopted diversified approaches to personalising dialogue systems. Li *et al.* [26] used reinforcement learning, processing user feedback for tailored responses. Zhang *et al.* [42] enhanced engagement by integrating user personas into the dialogue, while another work by Zhang *et al.* [41] utilised data from users' social media profiles for personalisation. Ghandeharioun *et al.* [12] focused on refining user experience by adapting to their language style and emotions. Despite these advances, challenges persist: the robustness of feedback-based reinforcement learning, the broader application of persona models, privacy concerns in social media-based insights, and the full implications of real-time personalisation methods, all require further exploration.

2.3.2 ***Style Adjustment in Dialogue Systems.*** Recent advancements in dialogue systems emphasise the importance of stylistic adjustments for enhanced user engagement. Wu *et al.* [40] proposed the "prototype-then-edit" methodology, highlighting the potential of deliberate stylistic modifications. Lample *et al.* [24] introduced an unsupervised style transfer model that maintains semantic content while altering style. Additionally, Keskar *et al.* [23] showcased refined stylistic control in large language models. Together, these contributions highlight the value of stylistic nuances in advancing AI-driven communication.

3 KEY FEATURES OF PERSONALISED CHAT ASSISTANT AND IMPLEMENTATION

Our user interface introduces an AI chat assistant designed to enhance user communication experiences. It has two primary functionalities as shown in Fig. 1.

Text Recommendation: This feature is useful when users are unsure about how to respond during conversations. Based on a user's preferences, the AI can suggest potential replies that align with a predetermined AI extraversion level and linguistic style. This grants users the flexibility to tailor the assistant's responses in terms of **1) Extroversion Level**: Three levels are available - introverted, average, and extroverted inspired by [38]. This ensures that the AI's recommended responses resonate with the user's desired interaction style. **2) Linguistic Style**: Users can set the AI to generate text with a specific tone, choosing between "humorous" and "respectful", according to results in the preliminary user survey of the appendix. This offers a personalised touch, making

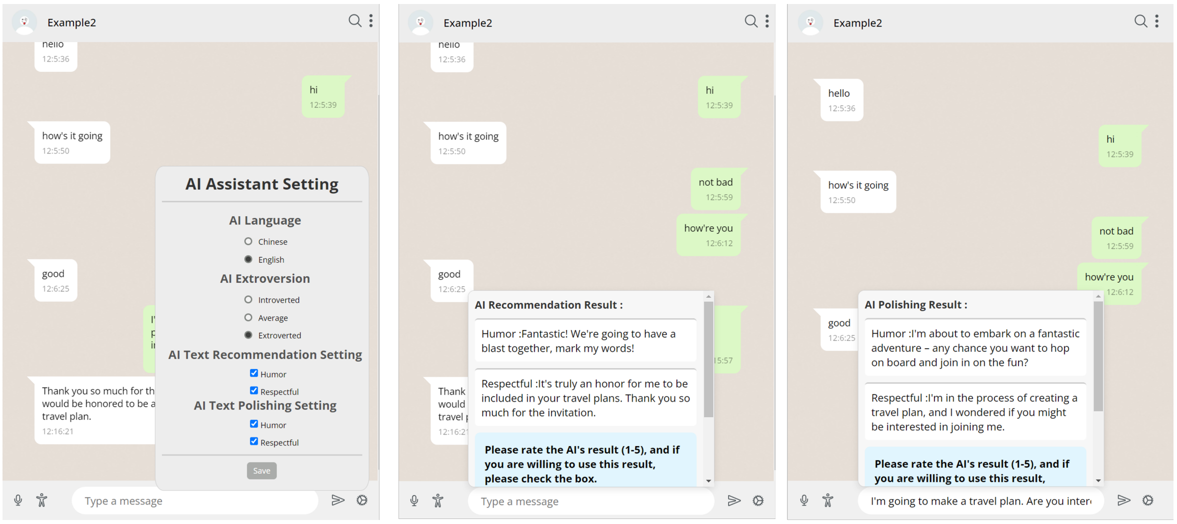

Figure 1: Interaction between two users of the AI chat assistant. The shown example is in English; however, all participants in the study consented to converse in Chinese.

the conversation more engaging or formal depending on the user's preferences and the perceived context.

Text Polishing: Beyond merely suggesting replies, the assistant can refine user-composed messages. Similar to the recommendation feature, the text polishing functionality is also customised based on the above two settings. This ensures clarity, coherence, and alignment with the chosen extroversion levels and linguistic styles.

Our Personalised AI Chat Assistant is underpinned by OpenAI's GPT-3.5 API[1]. In text recommendation mode, the system analyzes incoming chat content to generate contextually apt replies, according to user-specified stylistic preferences. In the text polishing mode, the system again analyses the incoming chat and processes the user's intended replies to produce refined text outputs. This adaptability, paired with an interface inspired by WhatsApp's web design and backed by an SQLite database for efficient data storage, led to a system suitable for the study of user-centric digital communication.

4 USER STUDY

4.1 Participants

We conducted an experiment involving a group of 28 Chinese participants. After a detailed explanation of the study's goals, these participants provided their informed consent by signing a form, thereby agreeing to participate in the experiment. Detailed information about each participant's characteristics is provided in Table 1 in supplementary material. The brief of the participants was that they would engage in conversations with people they had not met before, with a primary focus on the subject of travel. This setting was adopted to maintain consistency across all trials, in both the experimental condition of engaging with an unfamiliar person, and the topic of discussion.

[1]https://platform.openai.com/docs/api-reference

4.2 Study Setup

- **Participants Pairing**: We employed a random pairing strategy for our 28 participants, ensuring that participants in the control and experimental sessions did not interact with the same individuals, as shown in Table 1 in the supplementary material.
- **Experimental Process**: Each participant was involved in two distinct sessions as illustrated in Fig. 2. **Control Session:** First, participants engaged in a textual conversation with a person they were unfamiliar with, without the aid of the AI chat assistant. Upon conclusion, they rated their chat comfort level and their willingness to continue the conversation. **Experimental Session:** In the experimental phase, participants engaged in a distinct second text conversation with a new acquaintance, other than the person they engaged in the control session. In this session, the AI chat assistant was used, in the form of either AI-generated *recommendations* or by using the *text polishing* functionality. The limited range of supported language styles (two), and extraversion levels (three), allowed us to maintain a controlled linguistic environment, while offering the participants some meaningful choices, thereby catering to individual interaction preferences. Each session was timed to last approximately 20 minutes, ensuring uniformity in the duration of interactions across all trials. This structured approach was crucial for obtaining comparable data while allowing the AI's impact on the conversation dynamics to be observed effectively.

4.3 Metrics

System Usability Scale (SUS) serves as a robust instrument for evaluating the usability of systems such as AI chat assistants. It primarily assesses the ease and comfort level experienced by users during system interaction. An elevated SUS score is indicative of

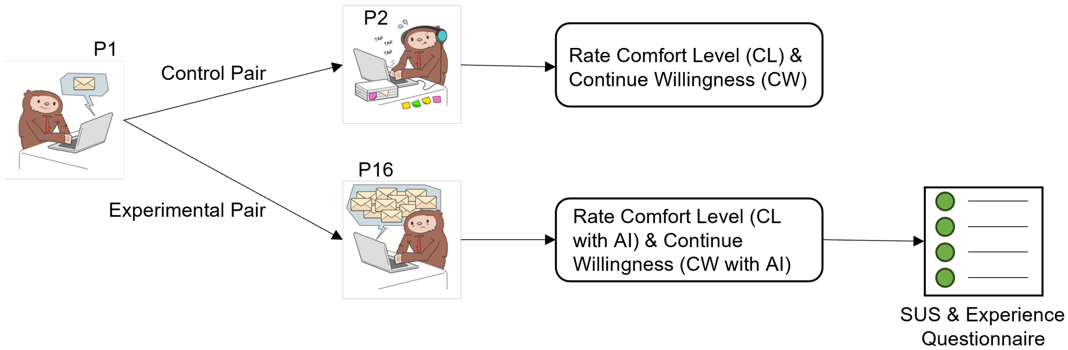

Figure 2: Two-phase experimental procedure: initial chat without AI followed by AI-assisted chat, with Comfort Level (CL), Continuation Willingness (CW) assessments, and post-experiment User Survey.

user-friendliness, implying that the AI assistant is perceived as more intuitive and manageable.

User Experience (UX) encapsulates the overall impression and response of users towards engaging with the AI chat assistant. This metric critically assesses the system's effectiveness, efficiency, and satisfaction quotient. A high UX score denotes that the AI assistant meets or exceeds user expectations in terms of functionality and responsiveness.

Comfort Level (CL) is used as a measure of the user's ease during initial textual interactions with new online contacts. Rated on a scale from 1 ("very uncomfortable") to 5 ("very comfortable"), with a neutral midpoint at 3, CL is instrumental in collecting user feedback for refining chat interfaces and enhancing the experience of online interactions [4].

Continuation Willingness (CW) quantifies a user's propensity to extend textual interactions following their initial engagement with new online contacts. Scored from 1 ("low willingness") to 5 ("high eagerness"), this metric is crucial for understanding user engagement and the probability of a subsequent sustained interaction in that digital environment, aiming at fostering a deeper connection [11].

5 RESULTS

5.1 Usability and User Experience

To understand users' perceptions of their interactions with our personalised AI chat assistant, we used the System Usability Scale (SUS). This measure offered insights into the system's perceived usability. Participants reported an average SUS score of 72.86 with a standard deviation of 8.04, which represents "good" usability [2]. Although the majority of scores clustered between 65 and 75, our Shapiro-Wilk test result (p = 0.0014) revealed that they did not follow a normal distribution. However, a median score of 70.00, an interquartile range [67.50, 75.63], and a bootstrap estimated 95% confidence interval for the median score [67.50, 73.75], indicate that most users found the AI chat assistant's usability to be satisfactory.

We employed the questionnaire in Schrepp *et al.* [31], to assess user experience across the six pivotal dimensions they proposed, by benchmarking our results against the standards they established. As illustrated in Fig. 3, the AI chat assistant excelled in the **Novelty** dimension, achieving a "Good" rating. This underscores users' perception of the AI chat assistant as a refreshing and innovative interaction tool. Furthermore, both **Perspicuity** and **Stimulation** received an "Above average" rating, highlighting the system's clear and intuitive design and its ability to invigorate user engagement.

However, the metrics for the remaining three dimensions fell within the "Below average" bracket, indicating areas with room for enhancement, including refining the aesthetics, streamlining the dynamic adjustments for increased efficiency, and bolstering the language model's robustness.

Additionally, the internal consistency of our evaluation, gauged via Cronbach's Alpha, revealed robust reliability for several dimensions. Specifically, **Attractiveness** (α = 0.77), **Dependability** (α = 0.76), and **Stimulation** (α = 0.80) all surpassed the empirically established threshold of 0.7. Other dimensions, namely **Perspicuity** (α = 0.65), **Efficiency** (α = 0.63), and **Novelty** (α = 0.65), hovered close to this benchmark, further supporting the credibility of our user experience findings.

5.2 Comfort Level and Continuation Willingness

Fig. 4 depicts the distribution of user ratings for comfort level and willingness to continue the conversation across the two experimental conditions. Descriptive statistics reveal a noticeable improvement in user responses when the AI chat assistant is active. Specifically, for comfort level (CL), we observed a mean score of $M = 2.86$ with a standard deviation of $SD = 0.69$ in the absence of AI, which increased to $M = 3.61, SD = 0.94$ with AI assistance, corresponding to a 26% improvement. For the continuation willingness metric (CW), the respective values were $M = 2.61, SD = 0.72$ without and $M = 3.79, SD = 0.82$ with AI, corresponding to an even bigger improvement of a 45.21%.

Normality tests with the Shapiro-Wilk method showed non-normal distributions for both CL ($p < 0.001$) and CW ($p < 0.001$). Given this, effect sizes were determined using Cliff's delta. The computed values indicated a medium effect for comfort level ($\delta_{\text{CL-AI}}$ = 0.43), suggesting a moderate improvement in comfort. In contrast,

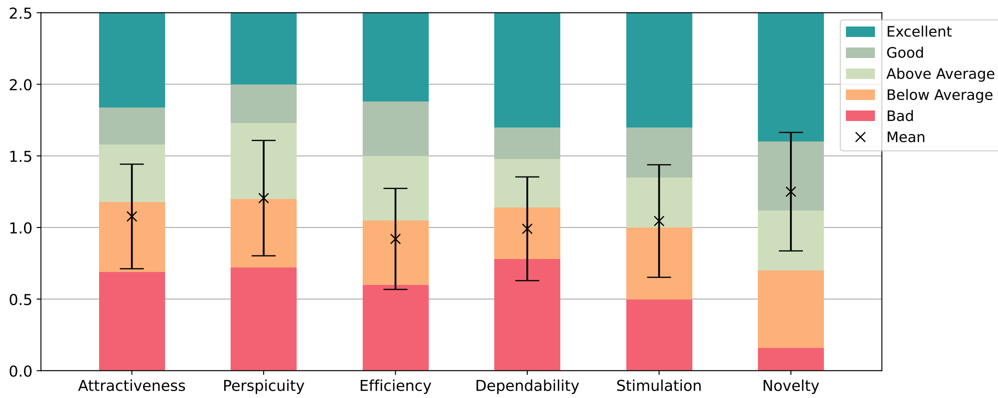

Figure 3: Distribution of scores across six key dimensions of user experience.

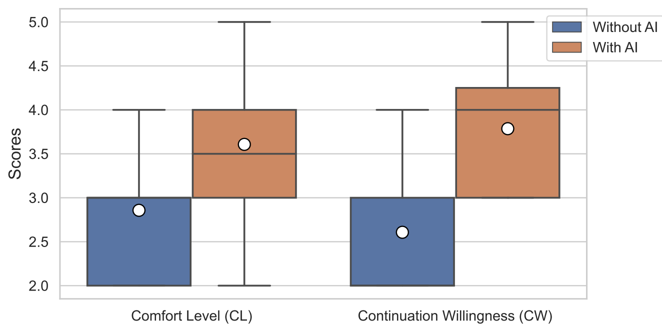

Figure 4: Boxplot outcomes of user Comfort Level (CL) and Continue Willingness (CW) in the presence vs. absence of AI intervention.

CW exhibited a large effect size ($\delta_{\text{CW-AI}} = 0.68$), indicating a significant increase in users' willingness to continue the conversation with AI assistance. The empirical findings underscore the positive effect of AI assistance in conversational interfaces. The enhanced comfort level suggests that users find the AI-enabled environment more conducive to interaction. Moreover, the pronounced increase in continuation willingness underscores greater user engagement and a generally positive user attitude towards conversational AI systems.

In summary, the integration of AI in chat systems evidently contributes to a more engaging and comfortable user experience. This finding holds a substantial promise for the design of future user-centric conversational agents and advocates for a human-centred approach in AI development.

5.3 Discussion

5.3.1 ***Usability and User Experience Reflection.*** Our findings reveal that the AI chat assistant significantly enhances text-based interactions in initial textual conversation scenarios. The system's high usability score signifies that users find it intuitive and user-friendly, a critical aspect in encouraging the adoption of new technology. The ease of use, minimal learning requirements, and straightforward interaction contribute to its favourable reception. In the realms of **Novelty** and **Perspicuity**, the AI chat assistant excelled, showcasing innovative characteristics and a clear and intuitive interface design. These attributes are essential in ensuring user engagement and satisfaction. However, the lower scores in other user experience dimensions, such as Attractiveness and Efficiency, point towards necessary enhancements. Addressing these areas could involve enhancing the visual appeal of the interface, expanding the range of dynamic adjustments for accommodating user preferences, and improving the sophistication of the language model. These improvements could significantly bolster the overall user experience, making the AI chat assistant not only a functional tool but also an aesthetically pleasing and efficient one.

5.3.2 ***Comfort Levels and Continuation Willingness.*** The improvement in users' comfort levels and willingness to continue interactions when assisted by the AI is particularly noteworthy. This aspect of our findings highlights the AI's role in creating a supportive environment for conversation, particularly in the initial, often awkward stages of unstructured interactions. The AI's ability to facilitate a more comfortable and engaging conversation environment is pivotal in contexts where establishing rapport and sustaining engagement are crucial.

5.3.3 ***Limitations and Future Directions.*** While our study offers some valuable insights, it is important to acknowledge its limitations. The focus on a specific ethnic demographic limits the generalisability of our findings, emphasising the need for future research to encompass a more diverse participant pool. Such an extension of our research could provide a more comprehensive understanding of the AI chat assistant's applicability across various cultural contexts. Variations in AI behaviour, even within the same extraversion level, could influence user perceptions. These variations, though minor, underscore the importance of consistency in AI interactions. Investigating these nuances can provide deeper insights into user-AI interaction dynamics. The potential over-reliance on the text polishing feature of the AI raises questions about long-term user dependency. Future studies should explore the implications of such reliance, particularly in terms of users’ communication skills development and their interaction with AI technology over time. Moreover, different users may interpret the same AI responses in various ways, indicating a need for more standardised measures or a broader range of qualitative data to capture the full spectrum of user experiences. In addition, the use of fixed AI personality profiles in our study presents a limitation. Future research should explore adaptable AI personalities that can dynamically adjust to match the user’s own personality, potentially enhancing the user experience and engagement further. This approach could address the needs of users with fluid personality traits, offering a more personalized and responsive interaction. Lastly, while our study sheds light on the potential of AI in enhancing initial text-based interactions, it also opens several avenues for future research on other types of human-human textual interactions. Exploring these will not only address the identified limitations but also expand our understanding of the complexities and potential of AI-assisted communication.

6 CONCLUSION

In the digital era, the generation of AI-powered, high-quality, text-based initial interactions is a very relevant yet understudied problem domain. Our results highlight the tangible benefits of the use of an AI chat assistant in elevating user experience. Users found the system highly usable and appreciated the heightened comfort it offered. Most notably, achieving consistent AI responses and accommodating diverse user personalities stand as promising areas for future refinement.

In summary, our research underscores the promising potential of AI in enhancing initial text-based exchanges. As technology advances, it is crucial to anchor developments in user-centricity. With sustained innovation, AI chat assistants have the potential to reshape digital communication, ensuring enriched and smooth human interactions.

REFERENCES

[1] David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics* 18, 2 (2014), 135–160.

[2] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[3] Naomi S Baron. 2004. See you online: Gender issues in college student use of instant messaging. *Journal of language and social psychology* 23, 4 (2004), 397–423.

[4] Susan Bergin and Ronan Reilly. 2005. The influence of motivation and comfort-level on learning to program. In *Proceedings of the 17th Workshop of the Psychology of Programming Interest Group*. Psychology of Programming Interest Group, Brighton, UK, 293–304. https://mural.maynoothuniversity.ie/8685/

[5] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. 2017. A survey on dialogue systems: Recent advances and new frontiers. *Acm Sigkdd Explorations Newsletter* 19, 2 (2017), 25–35.

[6] Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: user lifecycle and linguistic change in online communities. In *Proceedings of the 22nd International Conference on World Wide Web* (Rio de Janeiro, Brazil) *(WWW ’13)*. Association for Computing Machinery, New York, NY, USA, 307–318. https://doi.org/10.1145/2488388.2488416

[7] Marcel Danesi. 2017. *The semiotics of emoji: The rise of visual language in the age of the internet.* Bloomsbury Publishing, London, UK.

[8] Jacob Eisenstein. 2013. What to do about bad language on the internet. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Lucy Vanderwende, Hal Daumé III, and Katrin Kirchhoff (Eds.). Association for Computational Linguistics, Atlanta, Georgia, 359–369. https://aclanthology.org/N13-1037

[9] Jacob Eisenstein, Brendan O’Connor, Noah A. Smith, and Eric P. Xing. 2010. A latent variable model for geographic lexical variation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (Cambridge, Massachusetts) *(EMNLP ’10)*. Association for Computational Linguistics, USA, 1277–1287.

[10] Tira Nur Fitria. 2023. Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching* 12, 1 (March 2023), 44–58. https://doi.org/10.15294/elt.v12i1.64069 Number: 1.

[11] Mark Freiermuth and Douglas Jarrell. 2006. Willingness to communicate: can online chat help? 1. *International journal of applied linguistics* 16, 2 (2006), 189–212.

[12] Asma Ghandeharioun, Judy Hanwen Shen, Natasha Jaques, Craig Ferguson, Noah Jones, Agata Lapedriza, and Rosalind Picard. 2019. Approximating interactive human evaluation with self-play for open-domain dialog systems. *Advances in Neural Information Processing Systems* 32 (2019), 13658–13669.

[13] Kevin Gimpel, Nathan Schneider, Brendan O’Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan, and Noah A. Smith. 2011. Part-of-Speech Tagging for Twitter: Annotation, Features, and Experiments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Dekang Lin, Yuji Matsumoto, and Rada Mihalcea (Eds.). Association for Computational Linguistics, Portland, Oregon, USA, 42–47. https://aclanthology.org/P11-2008

[14] Bo Han and Timothy Baldwin. 2011. Lexical normalisation of short text messages: Makn sens a# twitter. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*. Association for Computational Linguistics, Portland, OR, USA, 368–378.

[15] Jeffrey T. Hancock, Catalina Toma, and Nicole Ellison. 2007. The truth about lying in online dating profiles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>San Jose</city>, <state>California</state>, <country>USA</country>, </conf-loc>) *(CHI ’07)*. Association for Computing Machinery, New York, NY, USA, 449–452. https://doi.org/10.1145/1240624.1240697

[16] Susan C Herring. 2007. A faceted classification scheme for computer-mediated discourse. *Language @ internet* 4, 1 (2007), 37 pages.

[17] Susan C. Herring. 2008. Language and the Internet. In *The International Encyclopedia of Communication*. John Wiley & Sons, Ltd, Hoboken, New Jersey, USA. https://doi.org/10.1002/9781405186407.wbiecl005

[18] Dirk Hovy and Anders Søgaard. 2015. Tagging Performance Correlates with Author Age. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Chengqing Zong and Michael Strube (Eds.). Association for Computational Linguistics, Beijing, China, 483–488. https://doi.org/10.3115/v1/P15-2079

[19] Zhiting Hu, Haoran Shi, Bowen Tan, Wentao Wang, Zichao Yang, Tiancheng Zhao, Junxian He, Lianhui Qin, Di Wang, Xuezhe Ma, Zhengzhong Liu, Xiaodan Liang, Wangrong Zhu, Devendra Singh Sachan, and Eric P. Xing. 2019. Texar: A Modularized, Versatile, and Extensible Toolkit for Text Generation. arXiv:1809.00794 [cs.CL]

[20] Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, and Eric P. Xing. 2017. Toward Controlled Generation of Text. In *Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 70)*, Doina Precup and Yee Whye Teh (Eds.). PMLR, Sydney, Australia, 1587–1596. https://proceedings.mlr.press/v70/hu17e.html

[21] Dell Hymes. 2001. *Foundations in sociolinguistics: An ethnographic approach.* Routledge, London, UK.

[22] Linda K Kaye, Helen J Wall, and Stephanie A Malone. 2016. “Turn that frown upside-down”: A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior* 60 (2016), 463–467.

[23] Nitish Shirish Keskar, Bryan McCann, Lav R. Varshney, Caiming Xiong, and Richard Socher. 2019. CTRL: A Conditional Transformer Language Model for Controllable Generation. https://doi.org/10.48550/arXiv.1909.05858 arXiv:1909.05858 [cs].

[24] Guillaume Lample, Sandeep Subramanian, Eric Smith, Ludovic Denoyer, Marc’Aurelio Ranzato, and Y-Lan Boureau. 2019. Multiple-Attribute Text Rewriting. In *International Conference on Learning Representations*. openreview.net, New Orleans, LA, USA, 20 pages.

[25] Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, and Bill Dolan. 2016. A Persona-Based Neural Conversation Model. https://doi.org/10.48550/arXiv.1603.06155 arXiv:1603.06155 [cs].

[26] Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016. Deep Reinforcement Learning for Dialogue Generation. https://doi.org/10.48550/arXiv.1606.01541 arXiv:1606.01541 [cs].

[27] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. “Blissfully Happy” or “Ready toFight”: Varying Interpretations of Emoji. *Proceedings of the International AAAI Conference on Web and Social Media* 10, 1 (2016), 259–268. https://doi.org/10.1609/icwsm.v10i1.14757 Number: 1.

[28] Shereen Oraby, Lena Reed, Shubhangi Tandon, T. S. Sharath, Stephanie Lukin, and Marilyn Walker. 2018. Controlling Personality-Based Stylistic Variation with Neural Natural Language Generators. https://doi.org/10.48550/arXiv.1805.08352 arXiv:1805.08352 [cs].

[29] Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov, and Alan W. Black. 2018. Style Transfer Through Back-Translation. https://doi.org/10.48550/arXiv.1804.09000 arXiv:1804.09000 [cs].

[30] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research* 21, 1 (2020), 5485–5551.

[31] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence* 4 (June 2017), 40–44. https://doi.org/10.9781/ijimai.2017.445

[32] Iulian V. Serban, Chinnadhurai Sankar, Mathieu Germain, Saizheng Zhang, Zhouhan Lin, Sandeep Subramanian, Taesup Kim, Michael Pieper, Sarath Chandar, Nan Rosemary Ke, Sai Rajeshwar, Alexandre de Brebisson, Jose M. R. Sotelo, Dendi Suhubdy, Vincent Michalski, Alexandre Nguyen, Joelle Pineau, and Yoshua Bengio. 2017. A Deep Reinforcement Learning Chatbot. arXiv:1709.02349 [cs.CL]

[33] Bayan Abu Shawar and Eric Atwell. 2007. Chatbots: are they really useful? *Journal for Language Technology and Computational Linguistics* 22, 1 (2007), 29–49.

[34] Crispin Thurlow. 2006. From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media. *Journal of computer-mediated communication* 11, 3 (2006), 667–701.

[35] Lisa Collins Tidwell and Joseph B Walther. 2002. Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human communication research* 28, 3 (2002), 317–348.

[36] G Vinodhini and RM Chandrasekaran. 2012. Sentiment analysis and opinion mining: a survey. *International Journal* 2, 6 (2012), 282–292.

[37] Tatiana A Vlahovic, Sam Roberts, and Robin Dunbar. 2012. Effects of duration and laughter on subjective happiness within different modes of communication. *Journal of Computer-Mediated Communication* 17, 4 (2012), 436–450.

[38] Sarah Theres Völkel, Ramona Schoedel, Lale Kaya, and Sven Mayer. 2022. User Perceptions of Extraversion in Chatbots after Repeated Use. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>New Orleans</city>, <state>LA</state>, <country>USA</country>, </conf-loc>) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 253, 18 pages. https://doi.org/10.1145/3491102.3502058

[39] J Wang, I Ivrissimtzis, Z Li, and L Shi. 2024. Impact of Personalised AI Chat Assistant on Mediated Human-Human Textual Conversations: Exploring Female-Male Differences. In *29th Annual ACM Conference on Intelligent User Interfaces*. Newcastle University, ACM Press, Greenville, SC, USA, 5 pages.

[40] Yu Wu, Furu Wei, Shaohan Huang, Yunli Wang, Zhoujun Li, and Ming Zhou. 2019. Response Generation by Context-Aware Prototype Editing. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 01 (July 2019), 7281–7288. https://doi.org/10.1609/aaai.v33i01.33017281

[41] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too? https://doi.org/10.48550/arXiv.1801.07243 arXiv:1801.07243 [cs].

[42] Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2020. DialoGPT: Large-Scale Generative Pre-training for Conversational Response Generation. https://doi.org/10.48550/arXiv.1911.00536 arXiv:1911.00536 [cs].

A SUPPLEMENTARY MATERIAL

B PRELIMINARY USER SURVEY

Prior to the commencement of an in-depth exploration into the domain of personalised AI chat assistants in initial text-based interactions, participants were duly informed of the study's aims and methodologies. Consent was obtained in alignment with the ethical guidelines of our institution. A preliminary survey was undertaken to ascertain the genuine interest and necessity for personalisation among potential users.

An online questionnaire[2] was formulated and disseminated among a heterogeneous group of individuals, encompassing an age spectrum of 15 to 44 years, along with other diverse demographic attributes. The survey witnessed participation from 162 individuals.

Table 1: Demographics and control group and experimental group pairs

Control Group Without AI (Age, Male/Female)	Experimental Group With AI
P1 (27, Female) & P2 (28, Male)	P1 & P16
P3 (28, Male) & P4 (28, Male)	P3 & P18
P5 (33, Female) & P6 (25, Female)	P5 & P20
P7 (27, Male) & P8 (28, Male)	P7 & P22
P9 (26, Male) & P10 (27, Male)	P9 & P24
P11 (28, Male) & P12 (27, Female)	P11 & P26
P13 (28, Female) & P14 (21, Female)	P13 & P28
P15 (27, Male) & P16 (29, Male)	P15 & P2
P17 (27, Male) & P18 (29, Female)	P17 & P4
P19 (27, Male) & P20 (27, Male)	P19 & P6
P21 (35, Female) & P22 (20, Male)	P21 & P8
P23 (35, Male) & P24 (33, Male)	P23 & P10
P25 (30, Male) & P26 (31, Female)	P25 & P12
P27 (26, Male) & P28 (28, Male)	P27 & P14

The questionnaire was initiated with queries aimed at discerning the general attitude towards chatbots and AI assistants, focusing on current levels of satisfaction and pinpointing areas necessitating enhancements.

A distinct section of the questionnaire presented participants with a range of language style options, such as 'formal', 'humorous', 'playful', 'serious', 'respectful', and 'offensive'. Participants were solicited to rate these styles, contemplating their ideal AI chat assistant, on a scale ranging from 1 (indicating "not needed at all") to 5 (signifying "highly needed"). Moreover, a rating bar was provided for participants to express their perceived need for a personalised AI chat assistant.

In the analysis of the results, scores exceeding 3 were considered indicative of a positive inclination. It was observed that a substantial majority of the respondents (82.67%) articulated a requirement for a more personalised AI chat experience. Regarding language style preferences, 'Respectful' and 'Humorous' were the predominant choices, preferred by 67.72% and 74.16% of respondents, respectively. Additionally, 'Playful' and 'Serious' styles were favoured by 66.93% and 53.54% of respondents, respectively. Conversely, a mere 33.86% of respondents exhibited a preference for an 'Offensive' tone.

This confirmation of user demand for personalisation, coupled with the identification of 'Respectful' and 'Humorous' as the most favoured language styles, facilitated the refinement of our subsequent experimental endeavours. This strategic realignment ensured that our study was deeply anchored in the actual preferences of users, thereby significantly enhancing its pertinence and potential for practical application.

[2]https://qfreeaccountssjc1.az1.qualtrics.com/jfe/form/SV_0VQb8OE6ZM5R5u6