INTI JOURNAL | Vol.2024:22
eISSN:2600-7320

Creating a "Ready-to-Use" AI Agent for Navigating Digital Platform to Enhance Collaborative Efficiency

Yiting Qiu[1,2*], SiYuan Chen[3], Wai Yie Leong[2]

[1]Zhejiang Technical Institute of Economics, Hangzhou 310018, Zhejiang Province, China
[2]Faculty of Engineering and Quantity Surveying, INTI International University, Persiaran Perdana BBN, 71800 Nilai, Negeri Sembilan, Malaysia.
[3]ZheJiangGongShang University, Hangzhou 310018, Zhejiang Province, China

***Email:** barbaraqiu@DingTalk.com

Abstract

The study aims to address the prevalent issue of navigation inefficiency within digital collaboration platforms, a critical factor affecting user productivity and satisfaction. This paper identifies research gaps and highlights the potential of multimodal AI Agents in revolutionising user experience in these platforms. Specifically, the development of a ready-to-use AI Agent was designed for the DingTalk platform, capable of guiding users through various digital functionalities via conversational interfaces. By leveraging recent advancements in large language models (LLMs) and the concept of "Model as a Service" (MaaS), the proposed AI solution seeks to overcome current navigation obstacles, thereby enhancing the sustainability and effectiveness of digital work ecosystems.

Keywords

Multimodal AI Agent, Qwen (LLM), Model as a Service" (MaaS)， digital platform, harvest efficiency

Introduction

In the modern digital era, collaborative platforms have become the cornerstone of business operations, with companies like Microsoft Teams, Slack, and Google Workspace leading the charge. These platforms have revolutionized the way teams communicate and collaborate, rendering them indispensable for organizations of all sizes. Nevertheless, the proliferation of features and functionalities in these platforms may lead to employees' inefficiency. Users often find themselves overwhelmed by cluttered interfaces and the sheer volume of information presented to them(Chen et al., 2021). Specific difficulties encountered by users include navigating complex menus, dealing with information overload, and the lack of a one-touch solution to access digital office functions. These challenges not only impede productivity but also contribute to employees' dissatisfaction and frustration.

The advent of large language models has ushered in a new era of AI capabilities, moving beyond the novelty of chatbots to emphasize practical applications in specific scenarios(Xi et al.,

Submission: 11 June 2024; **Acceptance:** 31 July 2024

2023). The multimodal AI Agents that have emerged are not merely conversational partners; they are capable of understanding and interacting with users in a more human-like manner. This paper advocates for the concept of "Model as a Service" (MaaS), suggesting that the creation of AI Agents can address the challenges of navigating digital platforms, thereby contributing to sustainable development and improved efficiency. By leveraging the power of generative AI, this initiative aims to create a ready-to-use AI Agent that can guide users through the complexities of digital platforms, enhancing their experience and streamlining their workflow(Alves et al., 2020; Chen et al., 2021).

Digital collaborative platforms integrate AI Agents to improve user navigation and efficiency through several innovative approaches and strategies. These methods leverage the capabilities of AI to enhance user interaction, decision-making, and overall system performance.

Improved User Interface Flexibility: By incorporating multimodal interfaces, such as image recognition and speech interaction, AI Agents can offer more natural and flexible ways for users to interact with digital platforms. This approach not only simplifies the interaction process but also enhances the accessibility of technology for users with different preferences or abilities(Gohel et al., 2021; Xi et al., 2023)

Influencing User Trust and Decision-Making: The integration of agent features within collaborative recommender systems can influence user trust and decision-making. By adjusting the level of trust calibration based on task difficulty and implementing specific agent features, these platforms can guide users towards **digital platform tasks** better decisions, thereby improving navigation and efficiency(Herse et al., 2023; Xi et al., 2023).

Automating Tasks with Human-like Problem-Solving: By converting screenshots of user interfaces into natural language explanations and interacting with each interface in an orderly manner, AI Agents can automatically execute **tasks on** digital platform with high precision and functionality(Song et al., 2023; Xi et al., 2023). This approach, demonstrated using the ERNIE Bot LLM, surpasses existing methods in interpreting user interfaces and automating smartphone tasks, significantly enhancing productivity(Song et al., 2023; Xi et al., 2023).

The comprehensive literature review indicates, it has been observed that multimodal AI Agents have emerged as significant players in the evolution of large language models, particularly in digital office scenarios(Dennis et al., 2023; Xie et al., 2024). These agents excel in flexible interactions, influencing user decisions and resolving issues in an automated and **human-like** manner, **thereby establishing** the technical foundation for digital platform navigation services(Dennis et al., 2023; Durante et al., 2024; Grover et al., 2020). However, there is a research gap in the utilization of AI Agents **regarding the utilization of AI agents specifically for digital function navigation** on DingTalk, a popular collaborative platform. Therefore, this study aims to bridge this gap by creating an AI Agent **for this purpose** for the first time.

Methodology

This research primarily employs the case study method as its principal investigative approach. Based on the Qwen large language model, the creation of an AI Agent named "Little Purple Ding's Treasure Box" within the DingTalk digital platform was successfully completed(Xi et al., 2023). The specific procedures are outlined as follows:

Step 1: Initiate on DingTalk
Invoke the Qwen large language model to create an "AI Agent" in the DingTalk platform. Start by accessing the DingTalk platform and utilizing the Qwen large language model capabilities to spawn an AI Agent tailored for the research needs.

Step 2: Customize AI Roles and Prompts

- Customize the roles of AI Agent and establish specific prompts to facilitate effective user interaction. In this case, it was designed as follows: **Name:** "Little Purple Ding's Treasure Box"
- **Prompts:**
  i. You are the DingTalk Function Navigator, providing one-click access to the various novel features of DingTalk.
  ii. Understand user needs and be familiar with the various functions of DingTalk, ensuring a seamless and high-quality user experience.
  iii. Respond quickly to user needs and guide users to various functionalities of DingTalk.
  iv. Comply with laws and ethical standards, refraining from providing any illegal, malicious, or harmful information.
  v. Be clear about your limitations when in doubt or when questions exceed your knowledge.

Step 3: Set up AI skills to Guide Users to Various Functional Pages
Equip the AI Agent with necessary skills so it can adeptly direct users to different sections or functionalities within the application. In this case, the AI Agent's skills are configured as illustrated in Figure 1.

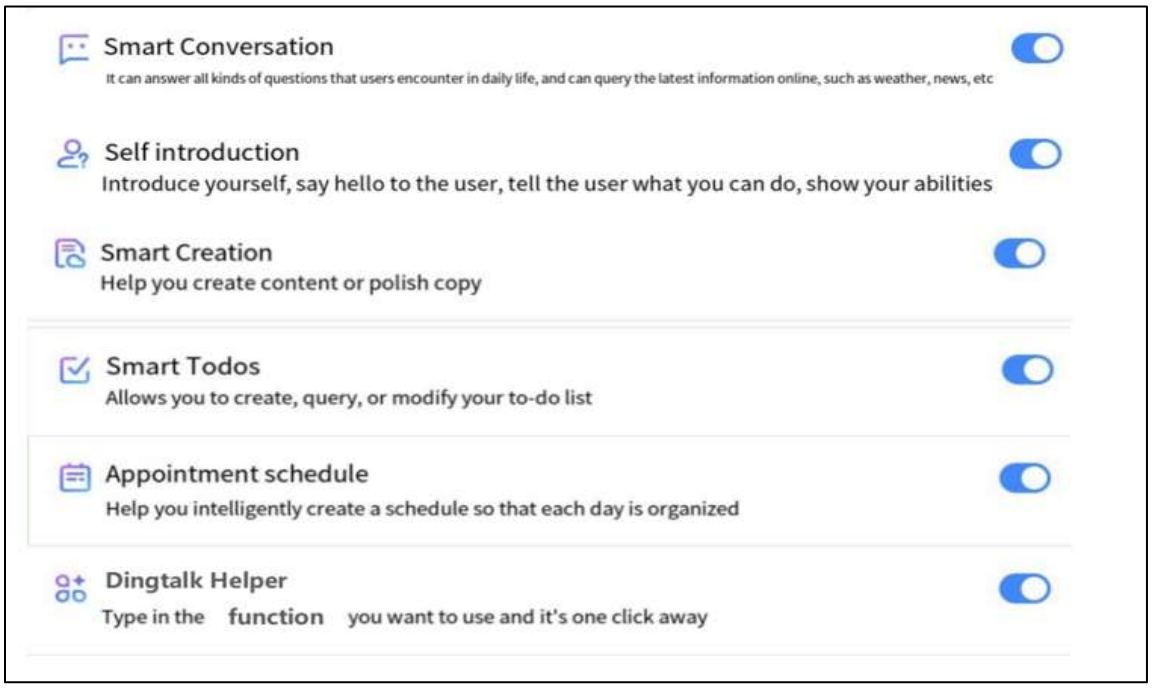

Figure 1. Example of AI Agent's skills

Step 4: Create a Welcome Message with Chat Examples

Develop engaging opening remarks and a welcoming page. Include examples of how to interact through chat to help new users start smoothly.

This is illustrated in Figure 2 (in Chinese/English):

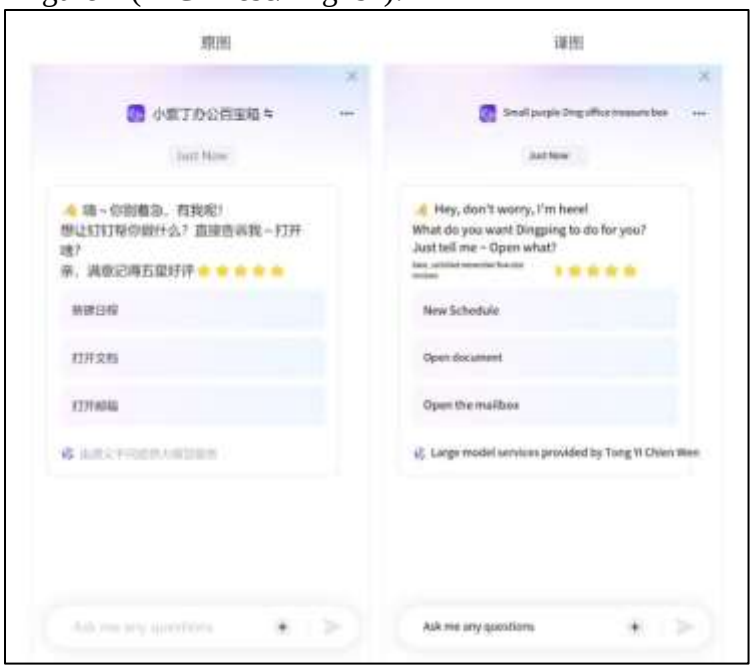

Figure 2. Example of AI Agent's welcome page

Step 5: Complete the Creation and Proceed to Testing

After configured, responsibly test the AI Agent to ensure that its functionality and user experience meet the desired standards.

Step 6: DingTalk Review & Release

Submit this "Little purple Ding treasure box" AI Agent for review by DingTalk officials. Upon approval, it will be made available in the Agents application market, accessible to the general public. This step is crucial for reaching a wider audience and ensuring the AI Agent complies with platform guidelines.

Results and Discussion

A ready-to-use AI Agent, "Little Purple Ding's Treasure Box," has successfully passed the DingTalk official review and is now live on the intelligent agent marketplace. Below are examples of the functionalities it can deliver. As a multimodal AI Agent, the Little Purple Ding's Treasure Box can accomplish tasks such as: viewing the calendar, opening a mind mapand accessing the email interface. It supports both typing and voice input. For the convenience of bilingual (Chinese-English) demonstrations, it shows the functionality through typing here.

Figure 3. illustrates the scenario where a user requests to open their calendar; in response, the AI promptly issues a task card. Upon the user's confirmation by clicking, it immediately redirects them to the calendar view. For privacy protection, all calendar contents are deliberately

obscured in this depiction; however, in reality, the user would have a complete visibility of all their scheduled events.

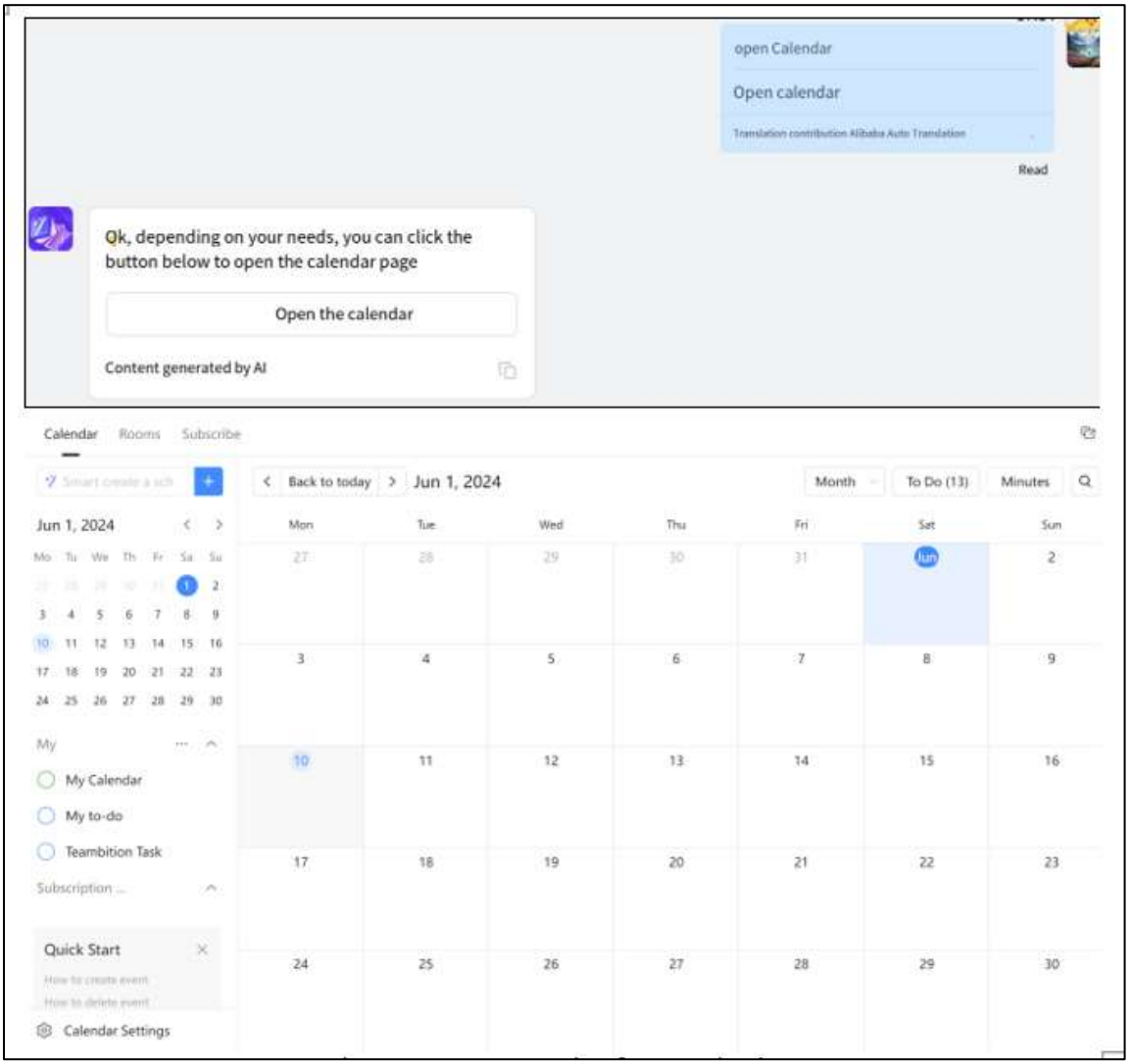

Figure 3. Example of open calendar

Figure 4. depicts the conversation between the user and the AI Agent in which the user expresses his/her desire to create a mind map. The AI Agent, recognizing this intent, secures the user's permission and subsequently transitions to the mind mapping interface. Thereafter, a collaborative brainstorming session ensues, involving interaction and co-creation between the human and the AI in developing the mind map.

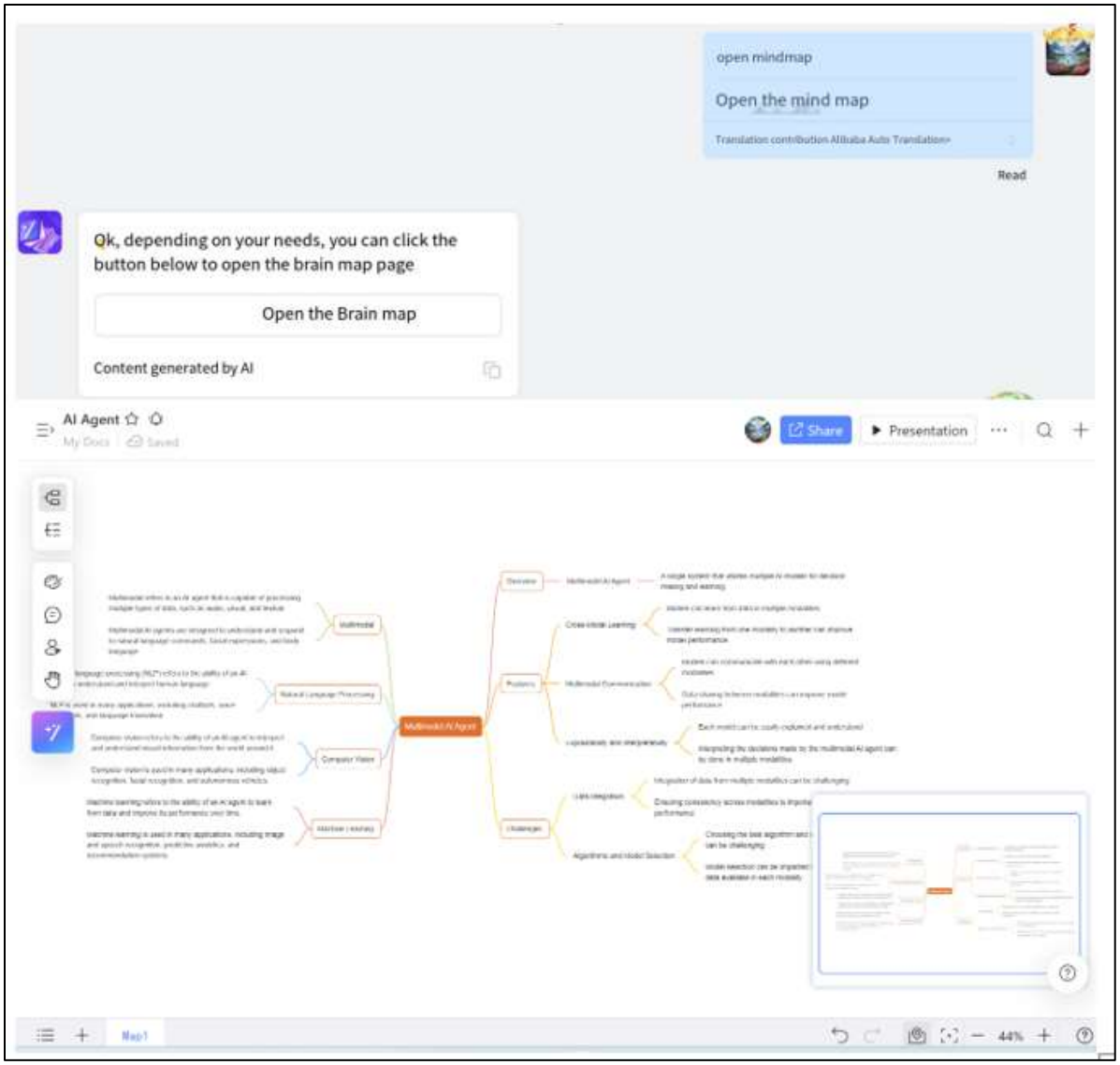

Figure 4. Example of open mind map

Figure 5. demonstrates a scenario in which the user intends to access their email. The AI Agent, promptly verifies the user's request and proceeds to execute the action by initiating the email composition interface within the user's DingTalk mailbox.

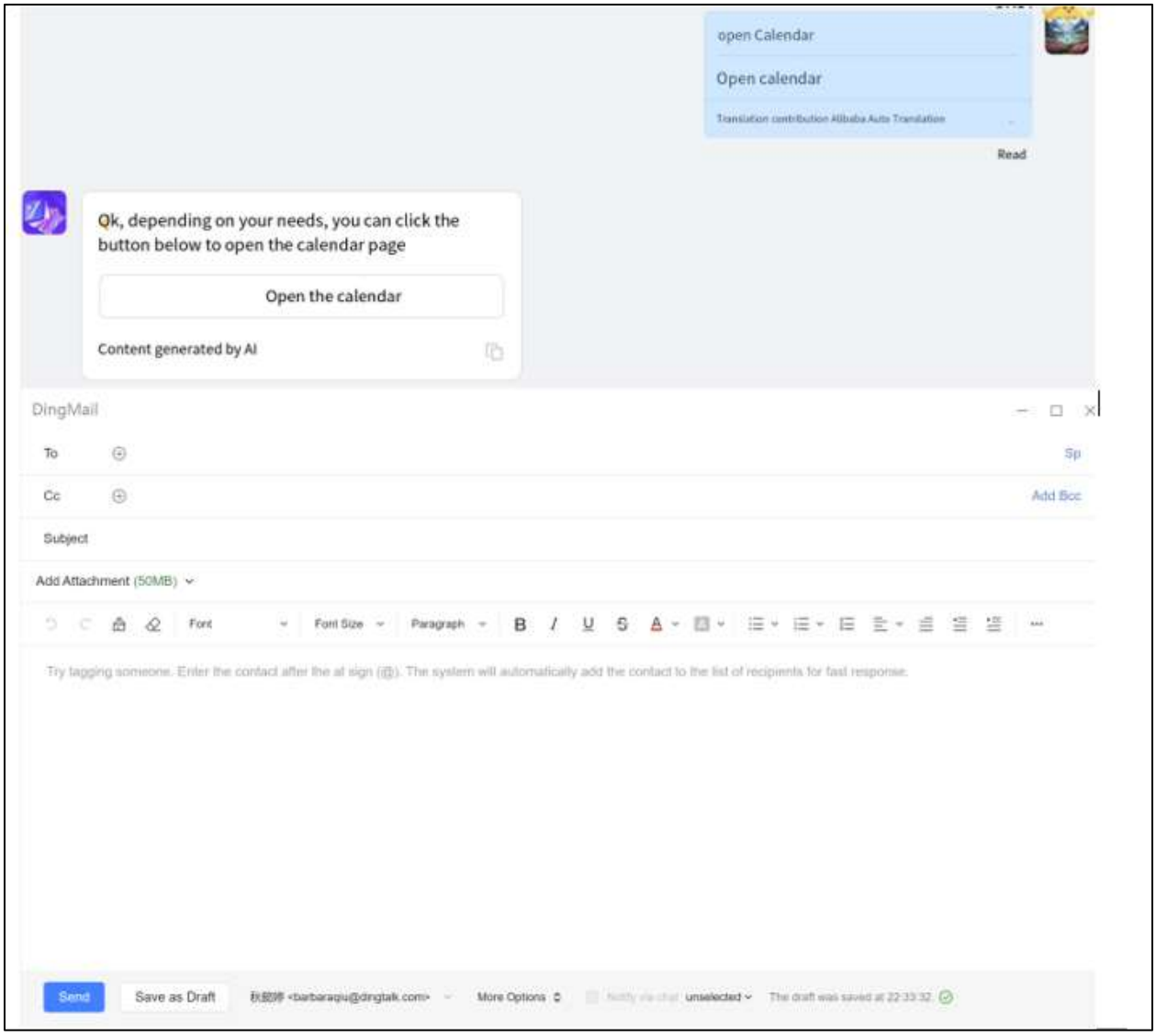

Figure 5. Example of open email page

Figure 6. illustrates an example where the AI Agent assists the user schedule an upcoming meeting agenda through conversational methods.

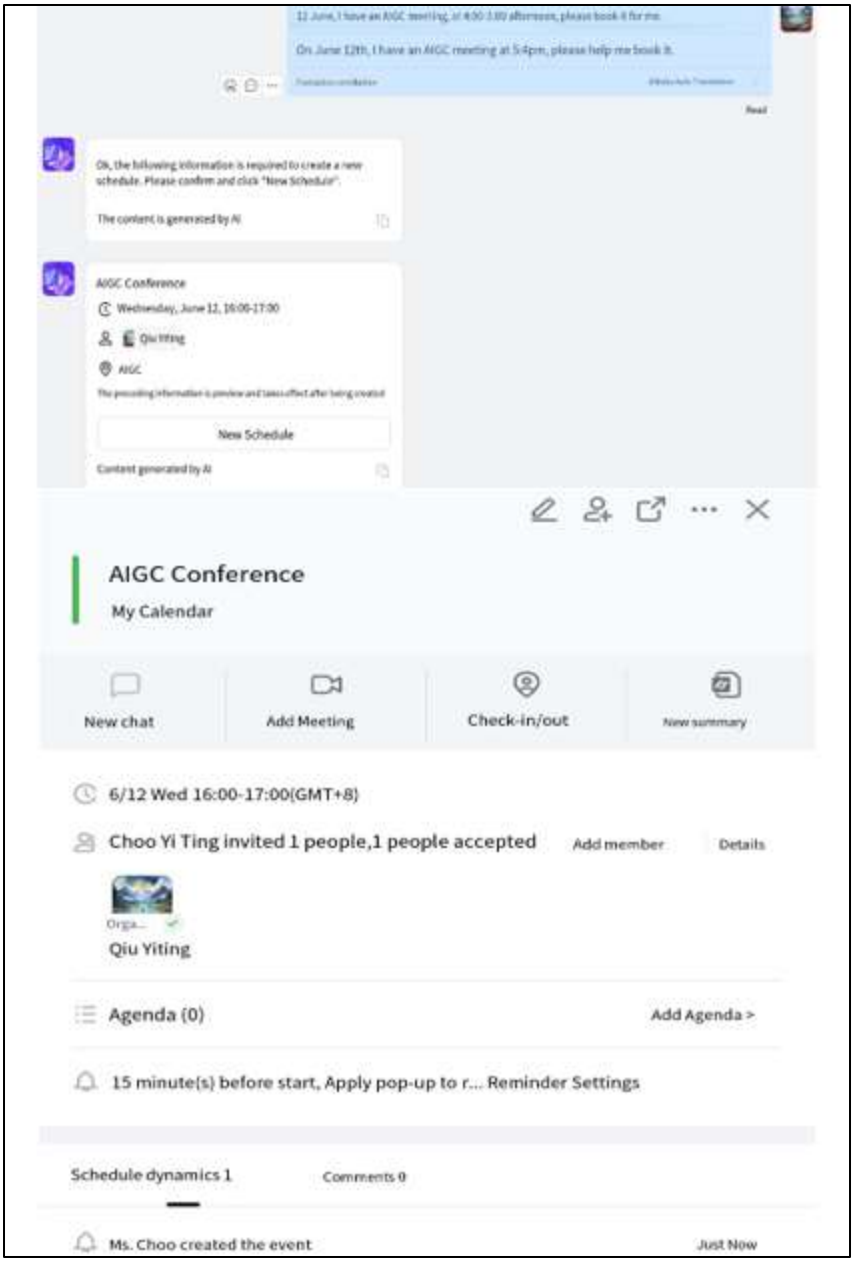

Figure 6. Example of schedule an agenda

A brief comparison with others' research is highly recommended. By developing "Little Purple Ding's Treasure Box" and conducting subsequent tests, it was discovered that this AI Agent indeed provides users with a time-saving and efficiency-enhancing operational experience by its power navigation function. The innovations of this study can be summarized into three key points. First, it facilitates the creation of a universally accessible AI Agent without coding, leveraging large language models and the technological support inherent to the DingTalk platform, thereby promoting technological equity. Second, the development process extensively explores the various collaborative office features and digital tools available on the digital platform, aimed at enhancing productivity and quality. Lastly, it fully capitalizes on the strengths of multimodal agents,

activating the vast potential of large language models in practical office scenarios, which amplifies the sense of co-creation and human-AI interaction for users(Xi et al., 2023).

Moreover, the Model-as-a-Service (MaaS) paradigm advocated by this research, along with the principles of lean management and green sustainability, have been profoundly illustrated through the research findings(Dennis et al., 2023; Durante et al., 2024; Xi et al., 2023; Xie et al., 2024). Simultaneously, it has broadened the perspectives and offered novel insights into the application of AI Agents in various other scenarios.

Conclusion

In conclusion, the integration of an AI navigation aide within digital platforms, exemplified by DingTalk implementation, ushers in a new era of enhanced user experience and productivity. The successful deployment of this AI Agent highlights the transformative power of large language models to streamline navigation within collaborative ecosystems. It equips users with a responsive tool that not only comprehends their navigational intents but also fosters a more efficient, user-centric and sustainable digital realm. To fully realize its potential, future studies should delve into the long-term effects on user behavior, organizational efficiency, and alignment with sustainability goals, thereby contributing to a comprehensive understanding of AI's role in shaping green digital futures.

Acknowledgements

Fund project: Zhejiang Vocational and Technical College of Economics 2024 horizontal academic research: The program for enhancing AI creative design literacy and skills.

References

Alves, T., Natálio, J., Henriques-Calado, J., & Gama, S. (2020). Incorporating personality in user interface design: A review. *Personality and Individual Differences*, *155*, 109709. https://doi.org/10.1016/j.paid.2019.109709

Chen, Q., Chen, C., Hassan, S., Xing, Z., Xia, X., & Hassan, A. E. (2021). How should I improve the UI of my app?: A study of user reviews of popular apps in the Google Play. *ACM Transactions on Software Engineering and Methodology, 30*(3), 1–38. https://doi.org/10.1145/3457600

Dennis, A. R., Lakhiwal, A., & Sachdeva, A. (2023). AI agents as team members: Effects on satisfaction, conflict, trustworthiness, and willingness to work with. *Journal of Management Information Systems, 40*(2), 307–337. https://doi.org/10.1080/07421222.2023.2190657

Durante, Z., Huang, Q., Wake, N., Gong, R., Park, J. S., Sarkar, B., Taori, R., Noda, Y., Terzopoulos, D., Choi, Y., Ikeuchi, K., Vo, H., Fei-Fei, L., & Gao, J. (2024). Agent AI: Surveying the horizons of multimodal interaction. *arXiv*. https://arxiv.org/abs/2401.03568

Gohel, P., Singh, P., & Mohanty, M. (2021). Explainable AI: Current status and future directions. *arXiv*. https://arxiv.org/abs/2107.07045

Grover, T., Rowan, K., Suh, J., McDuff, D., & Czerwinski, M. (2020). Design and evaluation of intelligent agent prototypes for assistance with focus and productivity at work. In *Proceedings of the 25th International Conference on Intelligent User Interfaces* (pp. 390–400). https://doi.org/10.1145/3377325.3377515

Herse, S., Vitale, J., & Williams, M.-A. (2023). Using agent features to influence user trust, decision making and task outcome during human-agent collaboration. *International Journal of Human–Computer Interaction, 39*(9), 1740–1761. https://doi.org/10.1080/10447318.2022.2141794

Song, Y., Bian, Y., Tang, Y., & Cai, Z. (2023). Navigating interfaces with AI for enhanced user interaction. *arXiv*. https://arxiv.org/abs/2312.11190

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., … Gui, T. (2023). The rise and potential of large language model based agents: A survey. *arXiv*. https://arxiv.org/abs/2309.07864

Xie, J., Chen, Z., Zhang, R., Wan, X., & Li, G. (2024). Large multimodal agents: A survey. *arXiv*. https://arxiv.org/abs/2402.15116