
OmniAI: A Multi-Agent Conversational Assistant for Task Automation

Priyanka Waghmare¹, Satchal Patil², Aditya Savale³, Prafull Satle⁴, Krishna Tilwane⁵

¹ Assistant Professor, Artificial Intelligence & Data Science Department, Ajeenkya DY Patil School of Engineering, Pune, India, priyanka.waghmare@dypic.in

^{2,3,4,5} Artificial Intelligence & Data Science Department, Ajeenkya DY Patil School of Engineering, Pune, India, satchalpatil04@gmail.com, adityasavale74@gmail.com, prafullsatle@gmail.com, krishnatilwane33@gmail.com

Abstract.

Business experts spend a good deal of their time on repetitive duties such as composing emails, preparing reports, running queries on structured data and making charts. There are a number of different types of automation tools, but many of them will need you to have programming skills or commands. These requirements restrict the amount of people who can use automated processes. This study proposes the multi-agent conversational process called OmniAI to give users a way to input into natural language commands that will be translated into digital action. The three main components of the OmniAI system are: intent detection, schema-based query construction and specialized agent systems for processing: emails, reports, visualizations and database queries. Experiments show it achieves an effectiveness of greater than 90 percent for carrying out tasks in each of the modules and has an average response time of less than two seconds, allowing useful perform. The suggested solution can be used in enterprise environments as well as educational institutions to reduce repetitive manual work and enhance efficiencies.

Keywords. Conversational AI; Natural Language Processing; Task Automation; Text-to-Sql; Multi-agent systems

1. INTRODUCTION

Professionals increasingly rely on digital tools to perform routine tasks such as writing emails, generating reports, querying structured data, and creating charts. These repetitive activities often require switching between multiple applications, increasing cognitive load and reducing productivity. Although automation technologies can streamline such workflows, many are designed primarily for technical users and require scripting, predefined commands, or familiarity with APIs. This limits their accessibility to non-technical users. Recent advances in natural language processing (NLP) and large language models (LLMs) provide new opportunities for human-computer interaction. By enabling computers to understand natural-language instructions, these technologies can automate complex digital workflows without requiring users to possess extensive technical knowledge.

2. RELATED WORK

Cognitive automation is moving towards more intelligent and adaptive enterprise systems, with a transitional phase between traditional Robotic Process Automation (RPA) [7] and cognitive automation. In recent research, tasks are delegated to autonomous agents in structured environments, known as agentic information systems [6]. The modern solutions use multi-agent orchestration frameworks which can facilitate coordination of special agents to perform complex tasks [1, 5]. These systems embed domain-specific functionality, such as using large language models to automate tasks [3] and multi-agent systems for data visualization [2]. Also, Text-to-SQL systems can benefit from the feedback obtained during execution to enhance queries precision [2]. The performance of the system is measured by accuracy and latency using a benchmarking system like TaskBench [4].

Existing traditional BI systems rely heavily on ETL-based data access, specialized interfaces, cloud-dependent infrastructure, and manual dashboard creation, limiting usability, privacy, and scalability. A gap remains in developing an integrated, locally deployable system that combines natural-language interaction, direct database access, feedback-driven accuracy, automated visualization, and collaborative reporting within a unified conversational framework.

3. SYSTEM ARCHITECTURE

OmniAI is organised in a five-layer hierarchical architecture, where each layer performs a specific part of the workflow execution pipeline. The layers are intended to be used in a sequential manner, where the output of one layer becomes the input to the other. A summary of all the layers and what they are composed of is given in table.

Table 1. OmniAI Architectural Layers and Component Responsibilities

Layer	Components	Function
User Interface	Chat, Voice, Dashboard	Accepts queries and displays results
Reasoning	NLP, Intent Router, Model Selector	Understands intent and selects models
Agents	SQL, Report, Email, Scheduling	Executes specific tasks
Data Processing	Database, Pandas, Plotly	Processes data and generates insights
Learning	Feedback Logger, Retraining	Improves performance from feedback

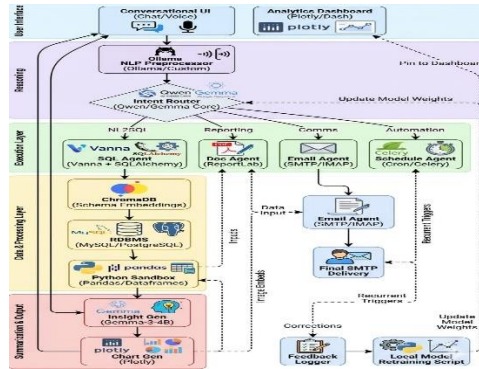


Figure 3.1. System Architecture Diagram

OmniAI processes natural-language queries through intent detection and specialized agents for SQL, reporting, and visualization, delivering results as charts, reports, or emails. Its key contributions include on-premises deployment, Qwen-Gemma multi-model routing, schema-aware NL2SQL with vector retrieval, and a unified conversational pipeline. Compared with LangChain and AutoGen, OmniAI achieves 92.5% query accuracy and reduces execution latency by 20–30% through specialized agent routing and schema-aware processing.

6. METHODOLOGY

It is a system with an agent, called a router, which is central and has knowledge about the user's intentions, and which decides which of the agents to send the task to. The NL2SQL agent converts your queries into SQL, the visualization agent generates graphs, the report agent generates structured output, and the email agent automates email communication.

7. RESULTS

The overall accuracy and success rates for the various tasks are high, in the range of 90% or more, and the reliability and performance of the system are consistently high. Automated emails achieve a 98.7% success rate. The system automates tasks through a natural-language interface with high accuracy and low response time. Its modular design provides flexibility, but performance depends on language-model accuracy and limited data testing. Future work will enhance contextual understanding and extend system capabilities.

Figure 7.1. Text to SQL Conversion Visualization

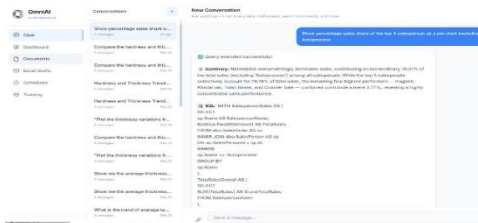
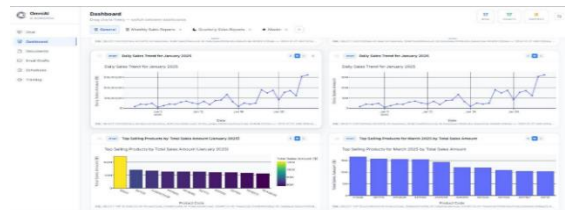


Figure 7.2. Dashboard Data



9. CONCLUSION

It proposes a multi-agent conversational system to automate digital tasks without the need of technical skills. Such a system has an accuracy of more than 90% in modules that respond within 2 seconds. Future work will focus on scalability, context and enterprise system integration.

10. REFERENCES

- [1] J. Hughes et al., “The Orchestration of Multi-Agent Systems: Architectures, Protocols, and Enterprise Adoption,” arXiv preprint arXiv:2601.13671, 2026. <https://arxiv.org/abs/2601.13671>
- [2] S. Fernando et al., “VizGen: Data Exploration and Visualization from Natural Language via a Multi-Agent AI Architecture,” arXiv preprint arXiv:2509.22218, 2025. <https://arxiv.org/abs/2509.22218>
- [3] R. Novelo, R. R. Silva, and J. Bernardino, “A Literature Review of Personalized Large Language Models for Email Generation and Automation,” Future Internet, vol. 17, no. 12, p. 536, 2025. <https://doi.org/10.3390/fi17120536>
- [4] Y. Shen et al., “TaskBench: Benchmarking Large Language Models for Task Automation,” arXiv preprint arXiv:2311.18760, 2024. <https://arxiv.org/abs/2311.18760>
- [5] W. Zhou et al., “Agents: An Open-source Framework for Autonomous Language Agents,” arXiv preprint arXiv:2309.07870, 2023. <https://arxiv.org/abs/2309.07870>
- [6] C. Engel, P. Ebel, and J. M. Leimeister, “Cognitive Automation,” Electronic Markets, vol. 32, no. 1, pp. 339-363, 2022. <https://doi.org/10.1007/s12525-021-00519-7>
- [7] P. Hofmann, C. Samp, and N. Urbach, “Robotic Process Automation,” Electronic Markets, vol. 30, no. 1, pp. 99–106, 2020. <https://doi.org/10.1007/s12525-019-00365-8>
- [8] H. Ye et al., “Cognitive Automation: A Survey on the Intersection of AI and Process Automation,” ACM Computing Surveys, vol. 56, no. 3, pp. 1–38, 2023. <https://doi.org/10.1145/3610214>

Biographies



Prof. Priyanka Waghmare is an Assistant Professor in the Department of Artificial Intelligence and Data Science Engineering at Ajeenkya DY Patil School of Engineering Pune. With nearly a decade of experience in academia, Prof. Priyanka has been instrumental in shaping the AI and Data Science curriculum and fostering a research-driven environment at ADYPOSE, Pune. Prof. Priyanka is currently pursuing a Ph.D. in Neural Networks.



Satchal Patil is currently pursuing a Bachelor of Engineering in Artificial Intelligence and Data Science at Ajeenkya DY Patil School of Engineering, Lohegaon, Pune. His areas of interest include artificial intelligence, machine learning, and data science.



Aditya Savale is currently pursuing a Bachelor of Engineering in Artificial Intelligence and Data Science at Ajeenkya DY Patil School of Engineering, Lohegaon, Pune. His areas of interest include artificial intelligence, machine learning, and data science.



Prafull Satle is currently pursuing a Bachelor of Engineering in Artificial Intelligence and Data Science at Ajeenkya DY Patil School of Engineering, Lohegaon, Pune. His areas of interest include artificial intelligence, machine learning, and data science.



Krishna Tilwane is currently pursuing a Bachelor of Engineering in Artificial Intelligence and Data Science at Ajeenkya DY Patil School of Engineering, Lohegaon, Pune. His areas of interest include artificial intelligence, machine learning, and data science.