
CAPSTONE PROJECT

Research Agent

Student Name : Mohammed Zubair A

College Name : Sir M. Visvesvaraya Institute of Technology

Department : Information Science and Engineering

OUTLINE

- **Problem Statement**
- **Technology used**
- **Wow factor**
- **End users**
- **Result**
- **Conclusion**
- **Git-hub Link**
- **Future scope**
- **IBM Certifications**

Problem Statement

The Challenge :

Researchers, students, and professionals often struggle to stay updated with the rapidly growing volume of academic publications, technical articles, and evolving research trends. Manually reviewing and synthesizing this information is time-consuming and inefficient.

Proposed Solution :

An AI Research Agent that uses Natural Language Processing (NLP) and Retrieval-Augmented Generation (RAG) to assist users in conducting efficient literature reviews, generating summaries, identifying research gaps, and recommending relevant papers.

Technology used

Technologies Used:

- **IBM Cloud Lite Services** : We utilized the free Lite tier for hands-on access to powerful AI services. This enabled cost-free development and deployment of our sophisticated agent for this capstone project.
- **IBM Granite Foundation Model** : This powerful, enterprise-grade LLM from IBM serves as the core intelligence of our agent. It was chosen for its strong reasoning and efficiency in generating accurate, relevant text from research queries.
- **Natural Language Processing (NLP)** : This core AI technology allows our agent to understand and interpret human language. It processes user questions and document text to enable meaningful analysis and interaction.
- **Retrieval-Augmented Generation (RAG)** : This advanced AI pattern allows the agent to first retrieve facts from a specific knowledge base. It then uses this retrieved information to generate more accurate and contextually relevant answers.

IBM cloud services used

IBM Cloud Services Used :

- **IBM Cloud watsonx.ai Studio** : This comprehensive, web-based platform served as our central development environment. It allowed us to access foundation models, manage project data, and build the AI agent from a single interface.
- **IBM Cloud Agent Lab** : This specific tool within watsonx.ai Studio provided the user-friendly interface to define our agent's instructions. We used it to connect the knowledge base, add tools like web search, and test agent behavior.
- **IBM Cloud watsonx.ai Runtime** : This service is the underlying computational engine that executes our agent in real-time. It provides the necessary processing power to run the Granite model and generate a response to the user's query.

Wow factors

This Agent Significantly reduces Research Time, Improves the Quality of literature reviews, and fosters Interdisciplinary Collaboration by making knowledge more Accessible and Actionable.

Unique Features:

- **Semantic Search** : Conducts searches across research papers, journals, and datasets.
- **Auto-Summarization** : Automatically summarizes selected papers to provide quick insights.
- **Smart Recommendations** : Recommends relevant research papers based on the user's current topic.
- **Hypothesis Generation** : Assists in creating simple hypotheses based on provided information.
- **Drafting Assistance** : Helps draft introductions and other sections of research papers.

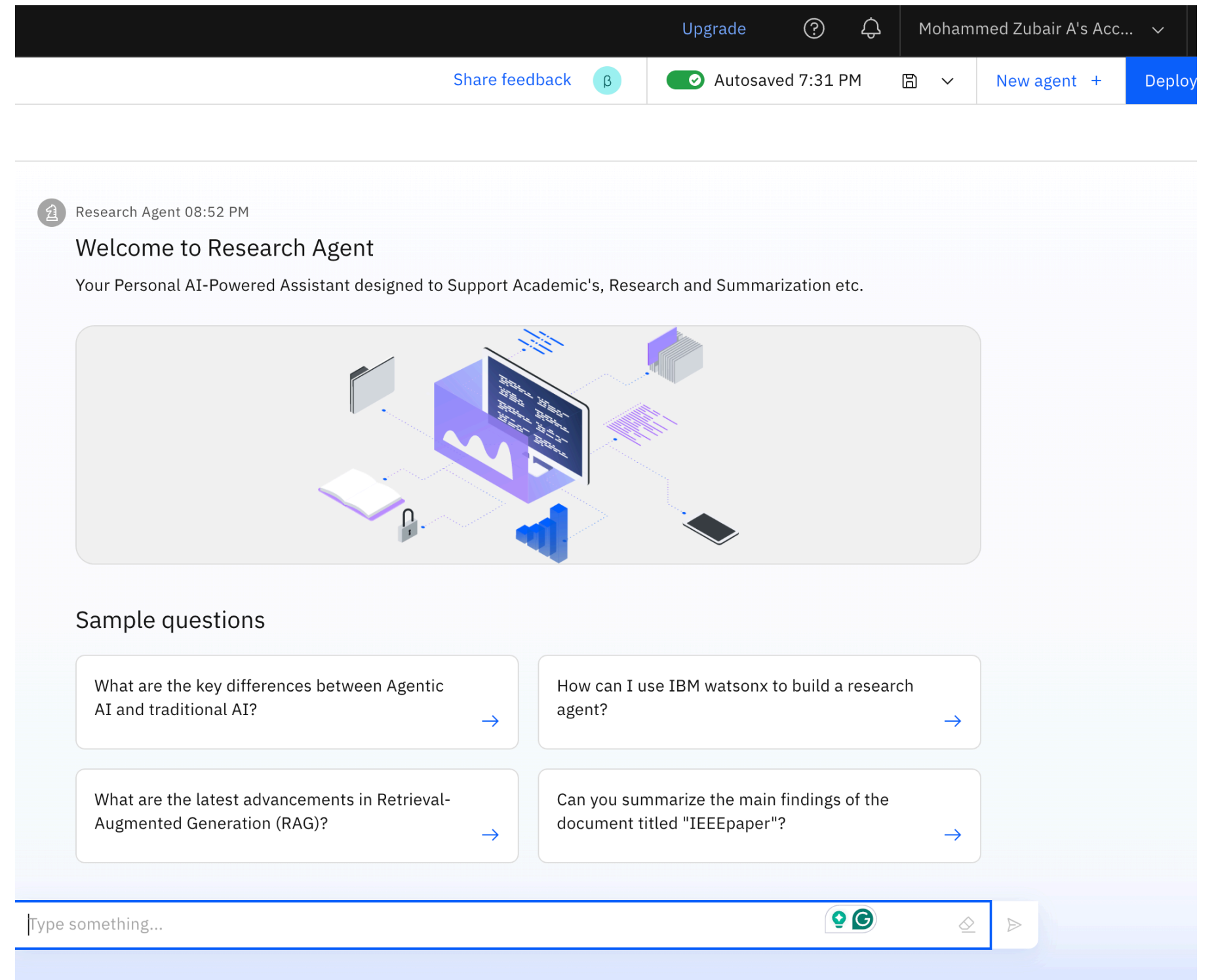
End users

- **Academic Researchers** : For literature reviews and staying current in their field.
- **Students**: To understand complex topics and find relevant sources for their assignments and theses.
- **Research Institutions and Universities** : To enhance the research capabilities of their staff and students.
- **Industry R&D Teams** : To accelerate innovation by quickly surveying the state-of-the-art in their domain.
- **Educators** : To prepare course materials and find relevant examples for teaching.

Results

Agent Welcome Screen

- **User-Centric Design** : This screen shows a clean, intuitive user interface, welcoming the user to the "Research Agent." The personalized description clarifies its purpose immediately.
- **Guided Interaction** : The "Sample questions" are a key feature. They guide the user on how to interact with the agent effectively, showcasing its primary capabilities like comparing AI types, summarizing documents, and explaining how it was built.
- **Ready for Input** : The prominent text input bar at the bottom indicates the agent is active and ready to receive user queries, providing a clear call to action.



Results

Answering a General Knowledge Question

- **Tool Usage Demonstration** : This image demonstrates the agent's ability to use its integrated tools (like a web search). The question about "advancements in RAG" is broad and not based on a specific uploaded document.
- **Information Synthesis** : The agent doesn't just return a link. It has retrieved information from multiple sources and synthesized it into a concise, numbered list, highlighting key advancements like multimodal retrieval and hybrid search.
- **Attribution and Transparency** : The "How did I get this answer?" link is crucial for building trust. It allows the user to see the sources and reasoning process the agent used, which is vital for academic and research applications.

Upgrade

?

🔔

Mohammed

Agent / Research Agent

Share feedback

β

Autosaved 7:31 PM

📁

▼

N

Agent preview

MZ

You: 07:31 PM

What are the latest advancements in Retrieval-Augmented Generation (RAG)?

📄

Research Agent 07:31 PM

Based on the search results, the latest advancements in Retrieval-Augmented Generation (RAG) include:

1. Multimodal retrieval, which enables the integration of text, images, and other data types.

2. Hybrid search, which combines different search mechanisms to improve retrieval efficiency.

3. Real-time retrieval, which allows for faster and more accurate retrieval of information.

4. Iterative LLM refinement, which improves retrieval outcomes through self-critique capabilities.

5. Speculative RAG, which enhances retrieval augmented generation through speculative techniques.

These advancements are expected to drive the future of RAG in 2025 and beyond, with potential applications in areas such as quantum computing, natural language processing, and computer vision.

► How did I get this answer?

🔄

Type something...

🗨️

🏠

Results

Summarizing a Specific Document

- **Core RAG Functionality** : This is the most important demonstration. The agent is answering a question about a specific document in its knowledge base ("IEEEpaper"). This shows it's not using general knowledge but is retrieving and analyzing the provided text.
- **Quantitative Data Extraction** : The agent successfully extracts specific, quantitative findings from the paper, such as "43% reduction in average coding time" and "sub-100ms latency." This proves its ability to go beyond simple text summarization and pull out key data points.
- **Contextual Accuracy** : The entire response is framed within the context of the document, demonstrating the power of RAG to provide highly relevant and factually grounded answers based on a custom knowledge source.

Upgrade

?

🔔

Mohammed Zubair A

Agent

Share feedback

β

Autosaved 7:30 PM

📁

▼

New agent

N

MZ

You: 07:30 PM

Can you summarize the main findings of the document titled "IEEEpaper"?

👤

Research Agent 07:30 PM

The main findings of the document titled "IEEEpaper" are:

- The collaboration protocol implements a modified operational transformation algorithm optimized for code editing scenarios, ensuring deterministic conflict resolution across all clients.
- The AI integration layer processes code in real-time, providing contextual suggestions based on the current document state and user input patterns.
- The system demonstrates a 43% reduction in average coding time and 67% improvement in error detection compared to traditional editors without AI assistance.
- The evaluation shows that the system maintains real-time synchronization with sub-100ms latency for up to 50 concurrent users, with a 99.7% success rate in automatic conflict resolution during simultaneous edits.
- The AI assistance module achieved 94.2% accuracy in code suggestions across multiple programming languages.
- The system performance remained stable under increasing user loads, with memory usage scaling linearly with active connections and CPU utilization remaining below 70% during peak loads.

Overall, the document presents a real-time collaborative programming environment with AI-powered assistance, demonstrating significant improvements in development productivity, collaboration efficiency, and code quality.

► How did I get this answer?

🔄

Type something...

🔍

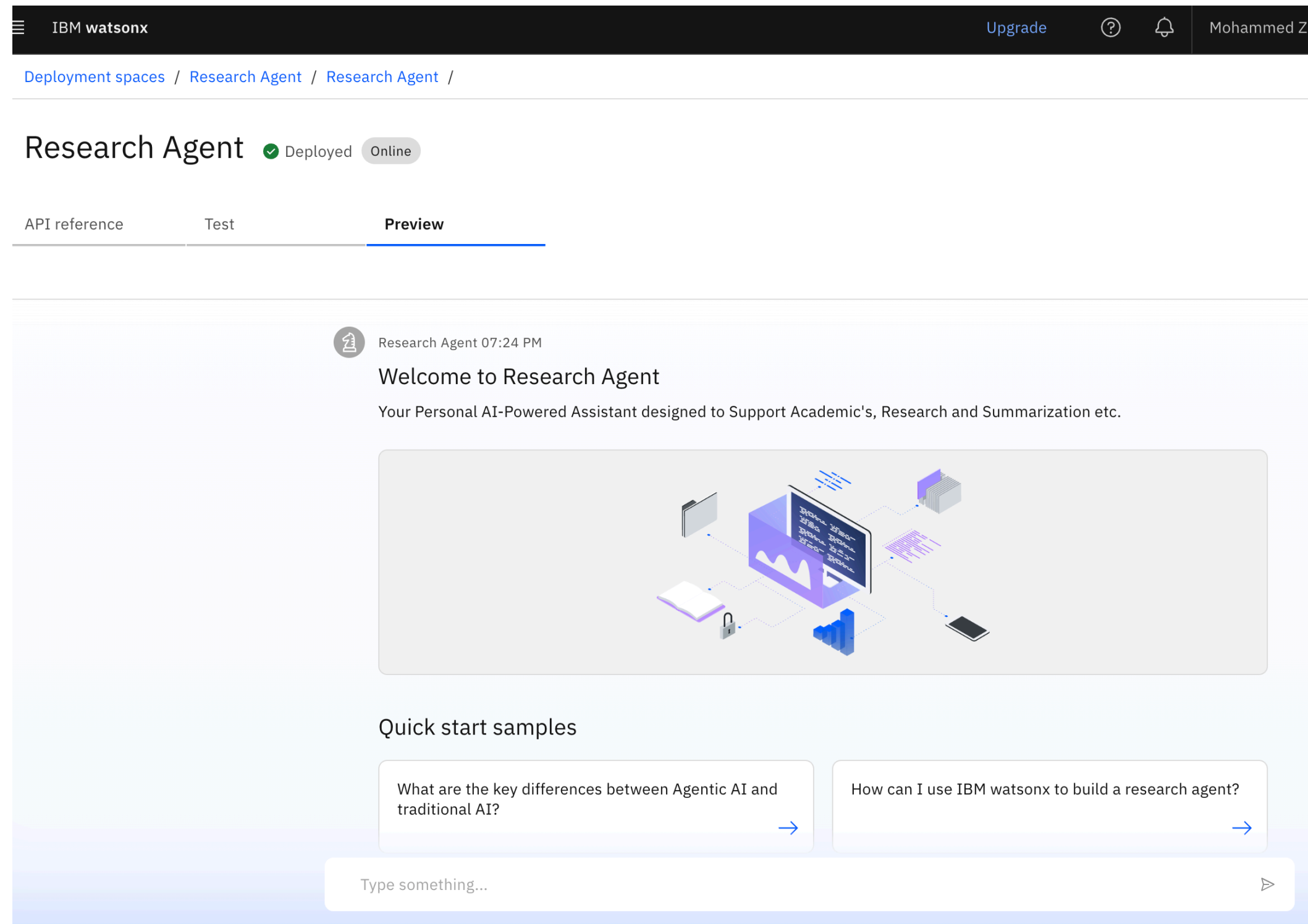
🗨️

➤

Results (Deployed)

Deployed Agent Welcome Screen

- **Successful Deployment** : The key takeaway from this image is the "Deployed" and "Online" status indicators at the top. This confirms that the agent has been successfully published as a live AI service on IBM Cloud.
- **Public-Facing Interface** : This screen shows the agent's preview interface from the "Deployment spaces." This is how an end-user or another application would interact with the live agent, separate from the internal build environment.
- **Consistency** : The UI remains consistent with the build preview, offering quick start samples and a clear input method, ensuring a seamless user experience from development to deployment.



Results (Deployed)

Interacting with the Deployed Agent

- **End-to-End Validation:** This image validates the entire project pipeline. A user is asking a question in the live, deployed interface, and the agent is responding correctly. This proves that the API endpoint is working and the backend service is functioning as expected.
- **Real-time Performance:** The agent provides a structured, detailed answer to the question about "Agentic AI vs. traditional AI." This demonstrates that the deployed service is responsive and can handle complex queries in real-time.
- **Scalability and Accessibility:** Having a deployed agent means it can now be integrated into other applications via its API or shared with other users. It has moved from a development prototype to a functional, accessible cloud service.

The screenshot displays the IBM Watsonx interface for a deployed agent. At the top, the header shows 'IBM watsonx' on the left, 'Upgrade' and a help icon in the center, and the user 'Mohammed Zu' on the right. Below the header, a breadcrumb trail reads 'Deployment spaces / Research Agent / Research Agent /'. The main section is titled 'Research Agent' with a green checkmark and 'Deployed' status, and a grey 'Online' button. Below this, there are three tabs: 'API reference', 'Test', and 'Preview', with 'Preview' being the active tab. The chat interface shows a user message at 07:24 PM: 'What are the key differences between Agentic AI and traditional AI?'. The agent's response, also at 07:24 PM, states: 'The key differences between Agentic AI and traditional AI are:' followed by a numbered list: 1. Autonomy: Agentic AI is designed to be more autonomous and adaptive, whereas traditional AI follows a script. 2. Adaptability: Agentic AI can interpret high-level goals, break them into sub-tasks, and adapt to changing conditions, whereas traditional AI is limited to executing pre-defined tasks. 3. Initiative: Agentic AI can take initiative to help achieve goals, whereas traditional AI requires explicit instructions. 4. Decision-making: Agentic AI introduces autonomy, adaptability, and proactivity in decision-making, whereas traditional AI relies on predefined parameters and explicit instructions. Below the list, a summary states: 'These differences enable Agentic AI to be more proactive, flexible, and effective in solving complex problems, whereas traditional AI is more limited in its capabilities.' At the bottom of the chat, there is a link '► How did I get this answer?' and a circular refresh icon. The bottom of the interface features a text input field with the placeholder 'Type something...' and a send button.

Conclusion

- The **AI Research Agent** successfully addresses the challenges of modern research.
- It saves valuable time by **Automating Repetitive tasks like literature searches and data extraction.**
- The AI Agent can **generate reports, suggest hypotheses, and even draft sections of research papers.**
- By leveraging IBM watsonx and the Granite model, this tool enhances **efficiency, accuracy, and innovation** in both academic and industrial R&D settings.

Future scope

- **Voice-Activated Research Assistant** : Implement voice commands for hands-free operation.
- **AI-Assisted Paper Drafting** : Enhance the agent's ability to help write, edit, and format entire research papers.
- **Research Gap Identification** : Develop capabilities to analyze the existing literature and identify novel research gaps.
- **Collaboration Mapping** : Suggest potential co-authors or institutions based on analyzing research interests and publication history.
- **Integration with Publishing Platforms** : Allow for direct submission or formatting for specific journals and publishing platforms.
- **Multilingual Research Support** : Add support for multiple languages to assist a global research community.

IBM Certifications

In recognition of the commitment to achieve
professional excellence



Mohammed Zubair A

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence



Issued on: Jul 15, 2025
Issued by: IBM SkillsBuild

Verify: <https://www.credly.com/badges/94248bea-3054-4b5a-b437-52f3124bc72b>



IBM Certifications

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to
Mohammed Zubair A

for the completion of

**Lab: Retrieval Augmented Generation with
LangChain**

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

GitHub Link

Github link : <https://github.com/Mdzub7/Research-Agent>

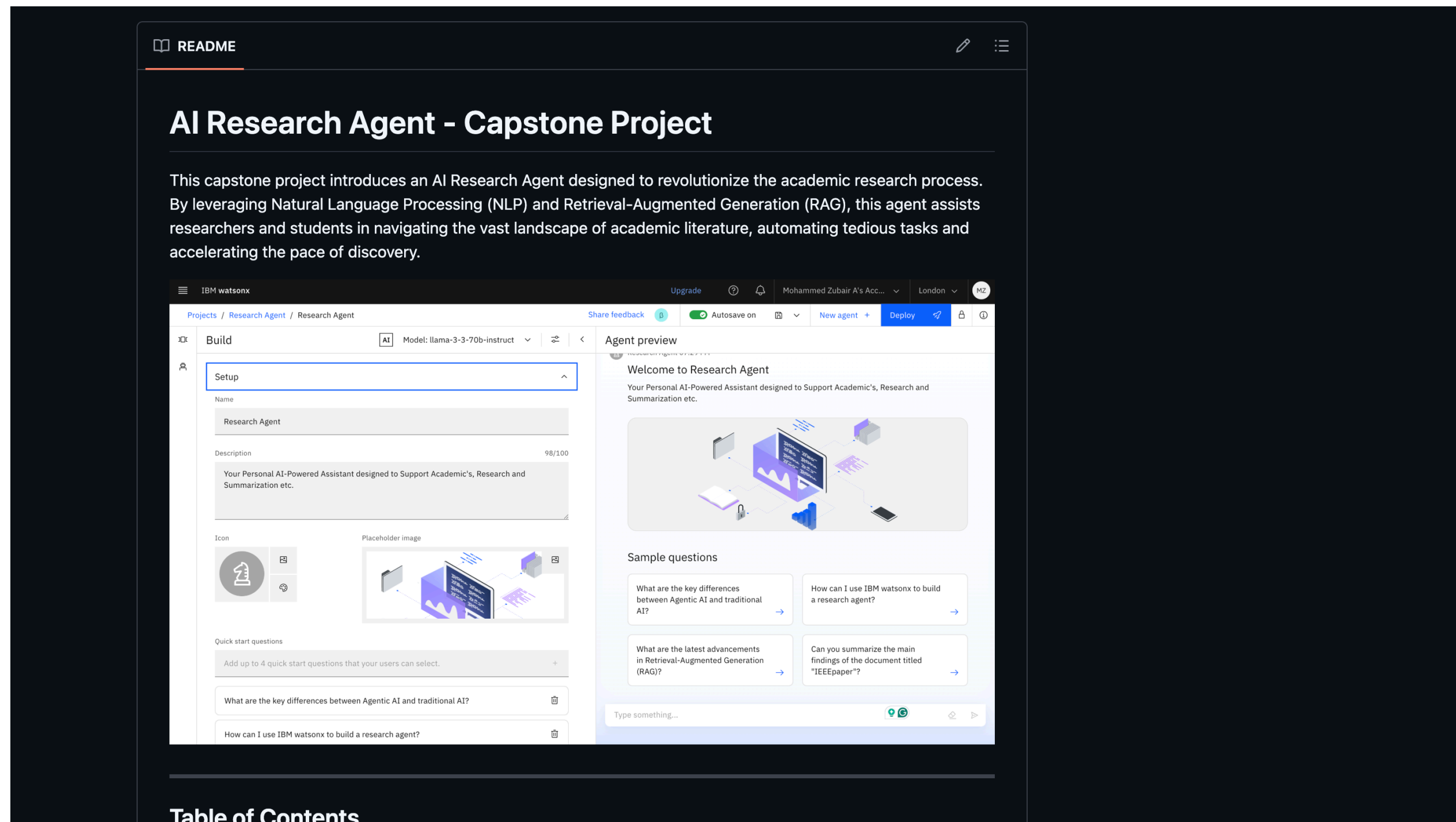


Table of Contents



THANK YOU