# CAPSTONE PROJECT

# RESEARCH AGENT

**Presented By:**
**Abinash Sharma**
**College Name- Maharaja Sriram Chandra Bhanja Deo University,Baripada, Mayurbhanj, Odisha**
**Department Of Computer Application**

edunet foundation

# OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

# PROBLEM STATEMENT

Research Agent The Challenge- A Research Agent is an AI system designed to assist with academic and scientific research tasks. It can autonomously search for literature, summarize papers, and organize references. Using natural language processing, it understands research questions and retrieves relevant information. The agent can generate reports, suggest hypotheses, and even draft sections of research papers. It saves time by automating repetitive tasks like citation management and data extraction. Research Agents enhance efficiency, accuracy, and innovation in both academic and industrial R&D. Technology - Use of IBM cloud lite services /IBM Granite is mandatory

# PROPOSED SOLUTION

The proposed system aims to create an AI-powered Research Agent capable of automating literature search, summarization, and reference management for academic and industrial researchers. This solution leverages **IBM Granite models** on **IBM Cloud Lite services** to ensure scalability, accuracy, and enterprise-grade reliability. The system will consist of the following components:

**Data Collection**

Collect academic papers and research articles from sources like IEEE, Springer, ArXiv, and PubMed.
Support multiple file formats (PDF, DOCX, URLs) for flexible research input.
Utilize real-time watsonx AI for updated research data.

**Data Preprocessing**

Extract and clean text from uploaded research documents.
Handle metadata (titles, abstracts, keywords) and remove irrelevant content (ads, formatting).
Convert data into embeddings for semantic search using IBM Granite embeddings.

# PROPOSED SOLUTION

**Natural Language Processing & AI Model**

Implement **Retrieval-Augmented Generation (RAG)** for precise literature search and contextual summarization.
Use IBM Granite models (large language models) for question answering and hypothesis suggestion.
Support multi-turn conversational research queries for better user interaction.

**Deployment**
Build a user-friendly web interface integrated with IBM Cloud Lite services.
Deploy the AI agent on **Watsonx AI** with Granite for real-time performance.

**Evaluation**
Measure summarization quality using  and accuracy of retrieved information using **precision-recall** metrics.
Gather researcher feedback to fine-tune the system continuously.
Conduct stress testing for performance and scalability on cloud infrastructure.

# SYSTEM APPROACH

**Strategy & Methodology:**
The Research Agent is developed using IBM Cloud Lite and IBM Granite to automate literature retrieval, summarization, and reference management. The approach ensures scalability, real-time performance, and academic-grade accuracy.

**System Requirements:**
**Hardware:** Cloud deployment via IBM Cloud Lite (lite tier for R&D).
**Software:** Python 3.x, IBM Watsonx services, Granite models.
**Storage:** Lightweight vector index (~1–5 GB).

**Libraries & Tools:**
**NLP & AI:** IBM Granite LLM.
**Deployment:** IBM Cloud Functions (Watsonx).

**Workflow:**
Summarizing & preprocessing (papers, abstracts).
semantic indexing.
Query handling with RAG (Granite model).
Cloud deployment & continuous evaluation.

# ALGORITHM & DEPLOYMENT

## Algorithm Selection

The Research Agent employs a **Retrieval-Augmented Generation (RAG)** architecture integrated with **IBM Granite large language models**. Watsonx ai is chosen because it combines semantic search (retrieving relevant documents) with generative AI (summarizing and answering queries), ensuring accurate and context-aware responses suitable for academic research.

## Data Input

Research papers and academic articles
Metadata: title, abstract, keywords
User queries (natural language research questions)

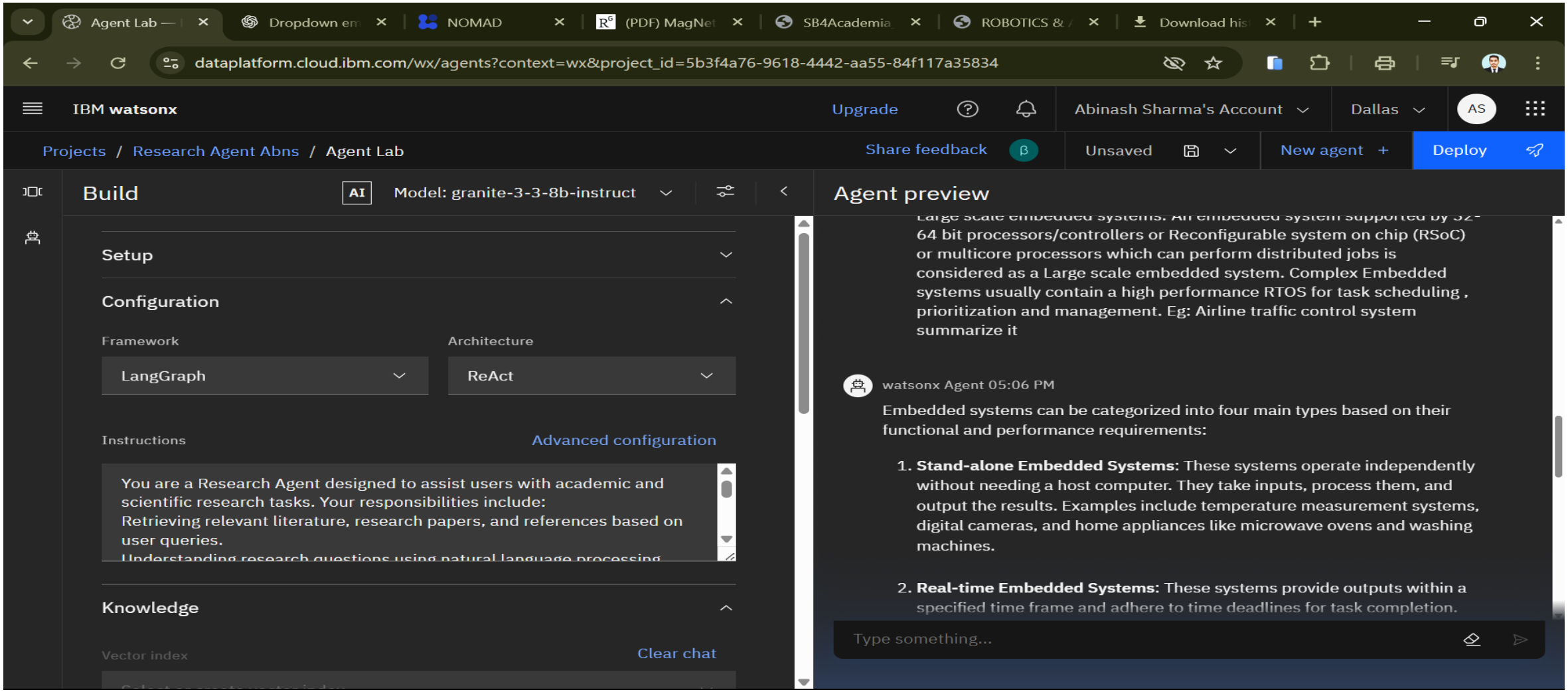
## Training Process

Preprocess documents and convert them into **vector embeddings** using IBM Granite embeddings..
Fine-tune the prompt templates for summarization and hypothesis generation to align with academic tone and accuracy.

## Prediction / Query Process

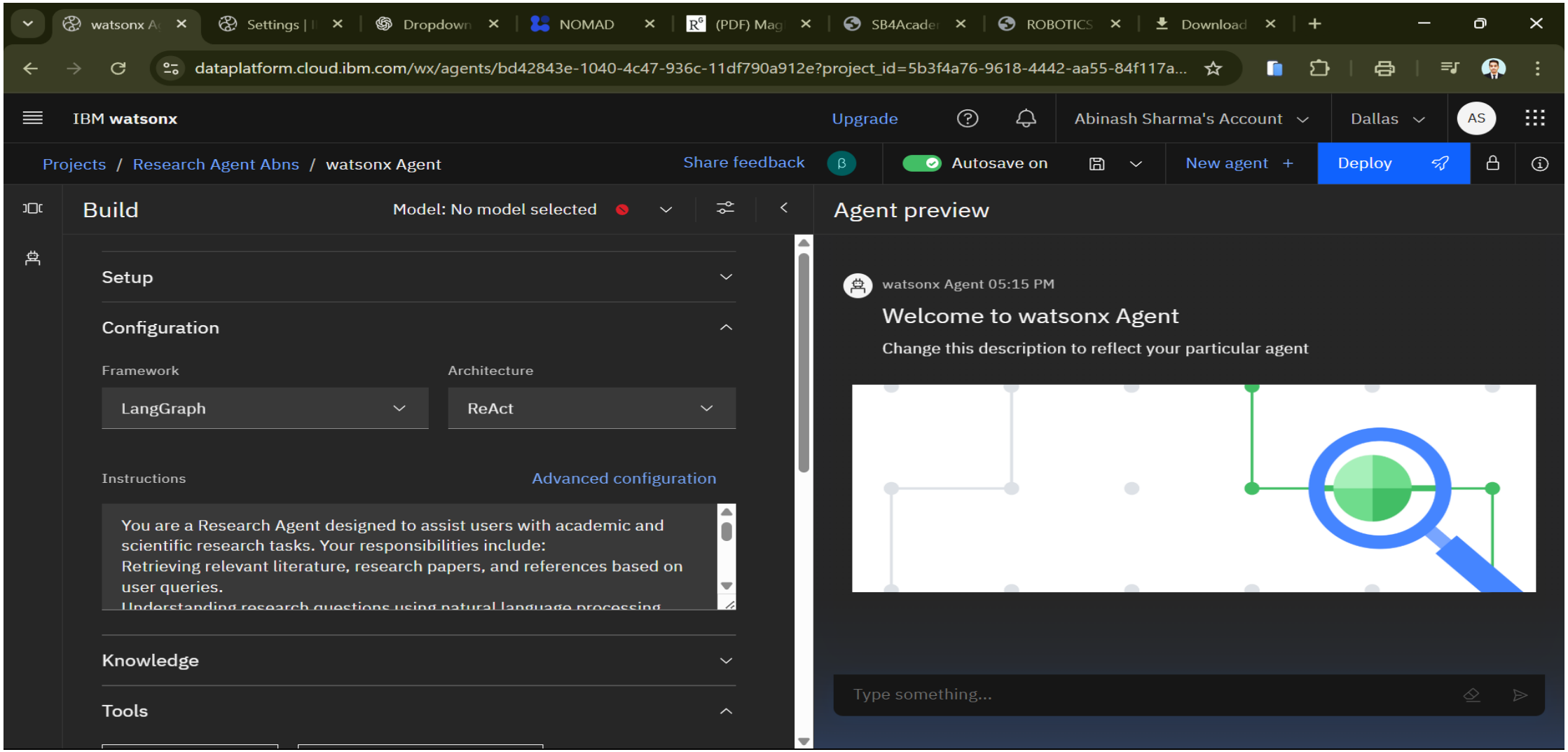
User submits a research question through the web interface.
Semantic search retrieves the most relevant documents from the vector index.
IBM Granite LLM generates a **concise, context-aware summary or answer**.
Results are displayed in structured format (summary, references, and possible hypotheses).

# RESULT



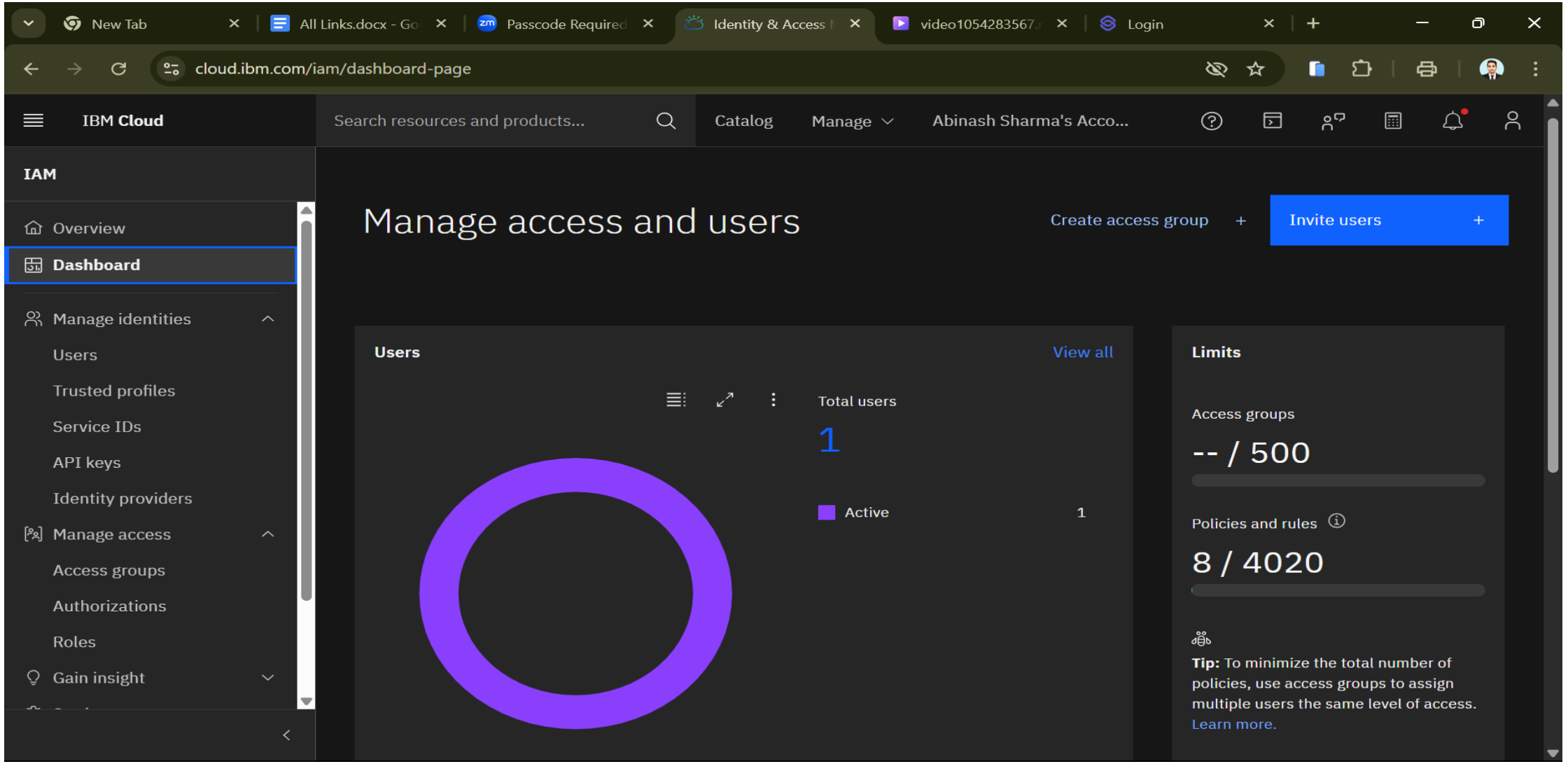

edunet foundation

# RESULT



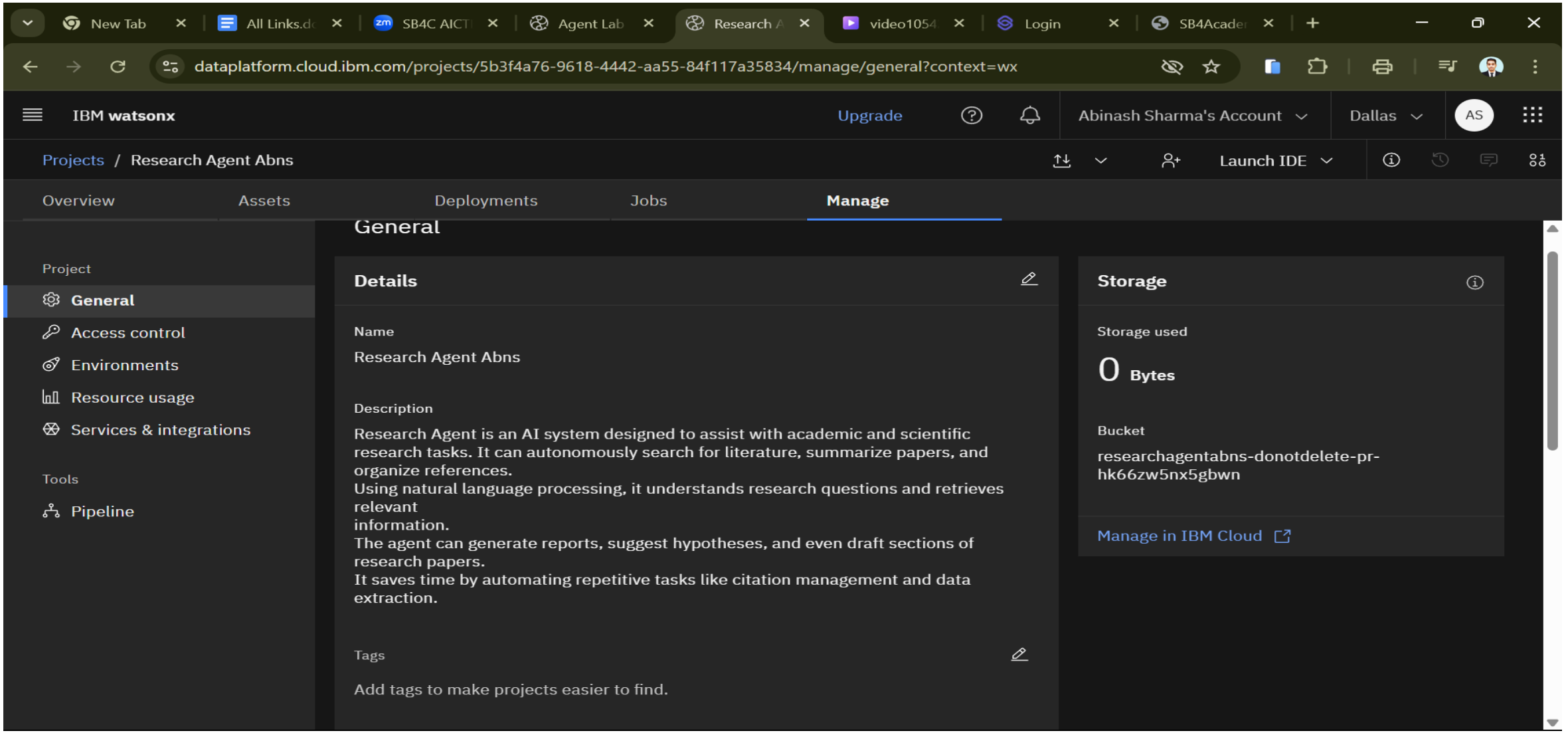

edunet foundation

# RESULT



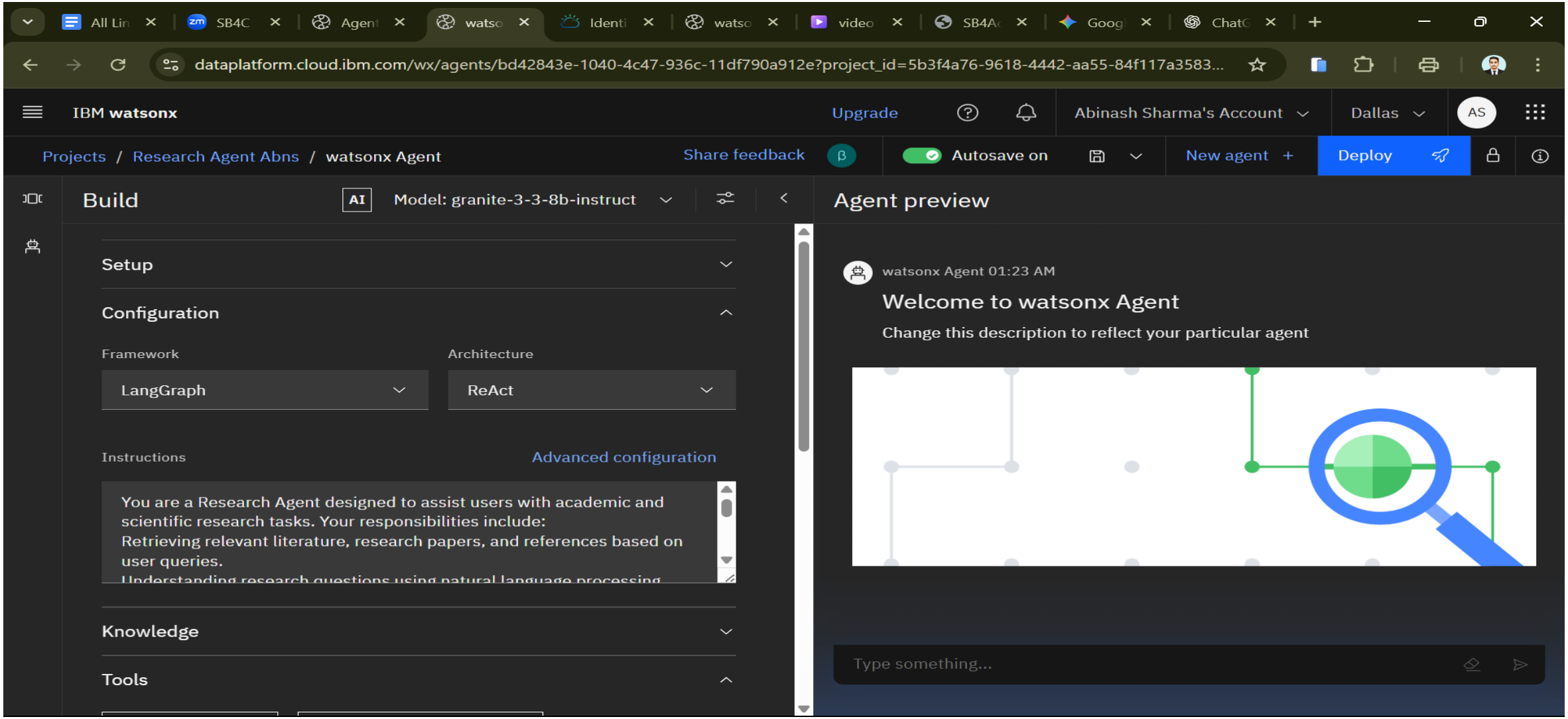

edunet foundation

# RESULT



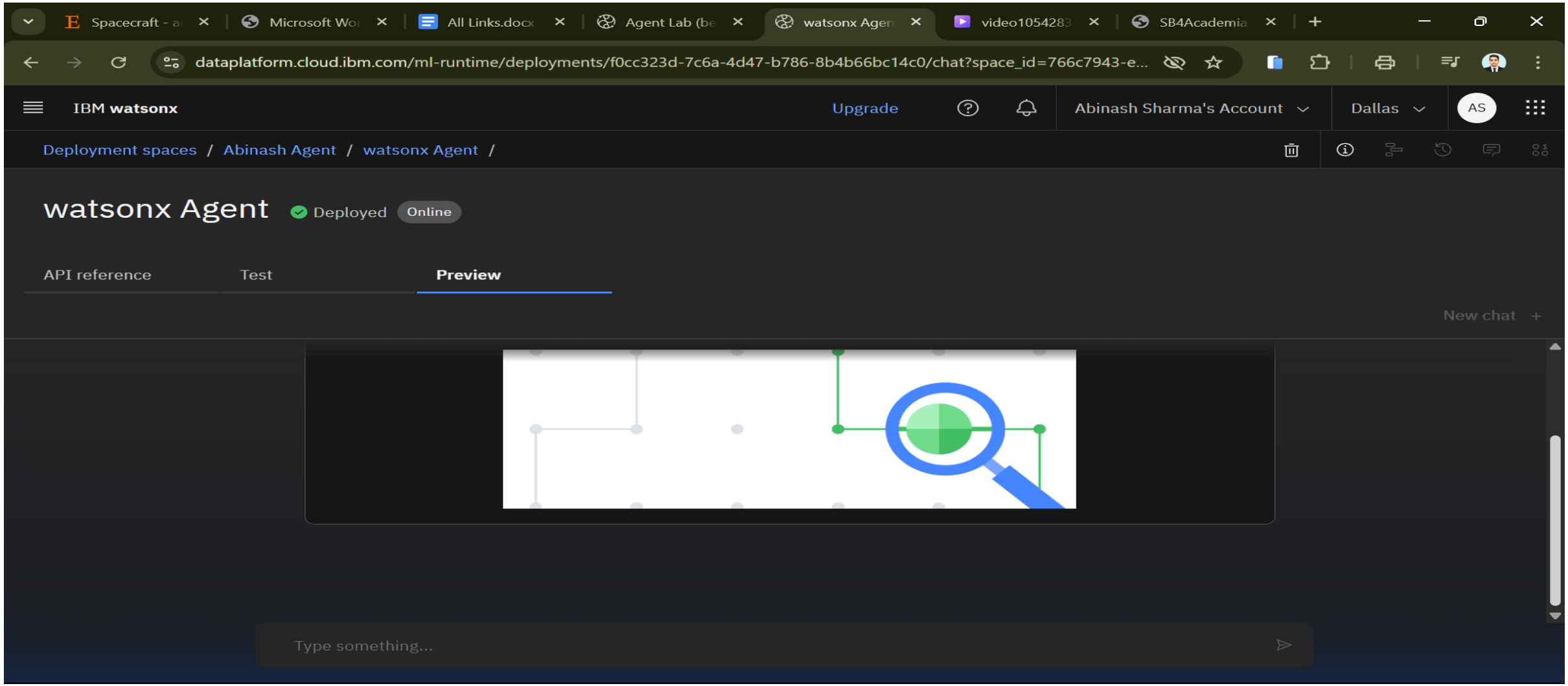

edunet foundation

# RESULT



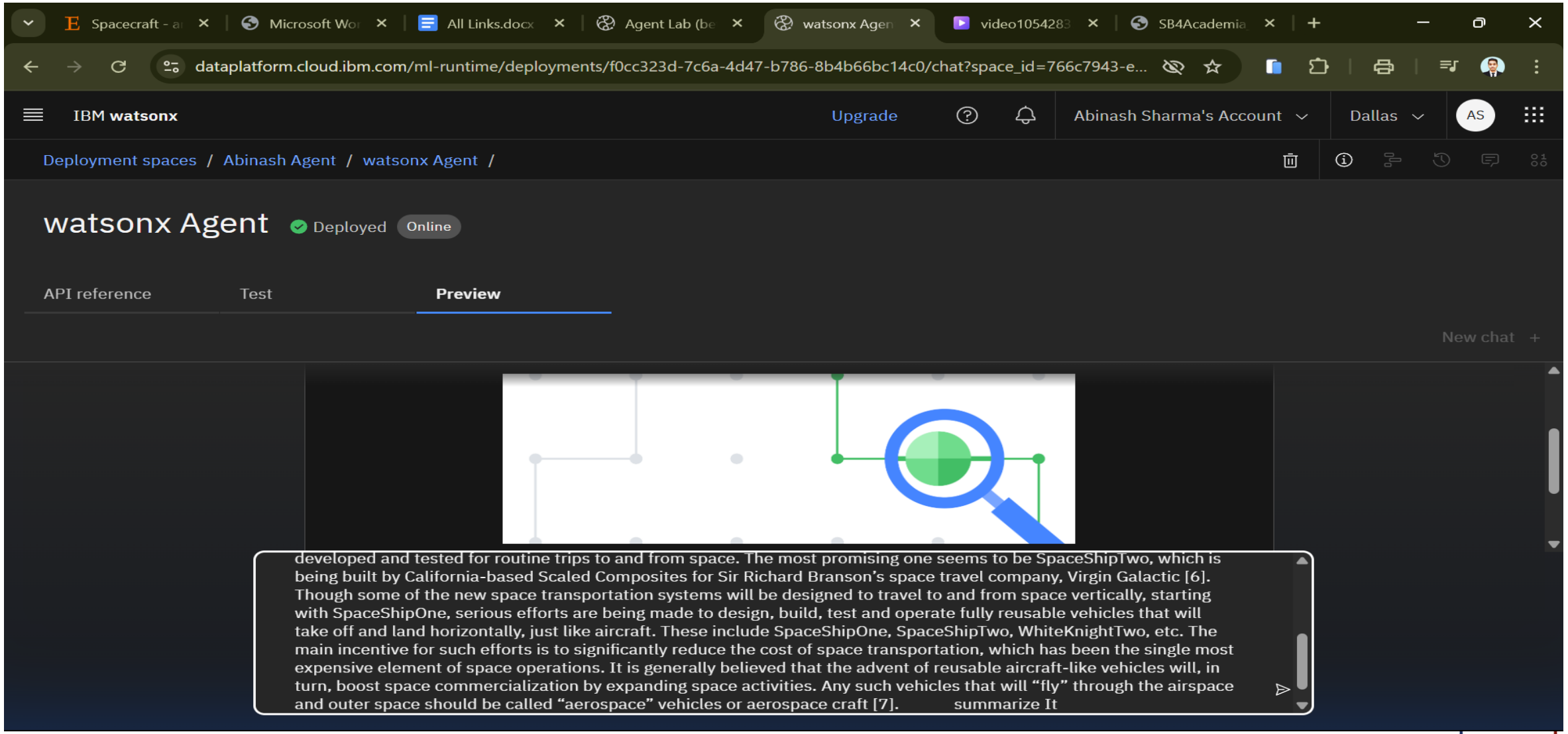

edunet foundation

# RESULT

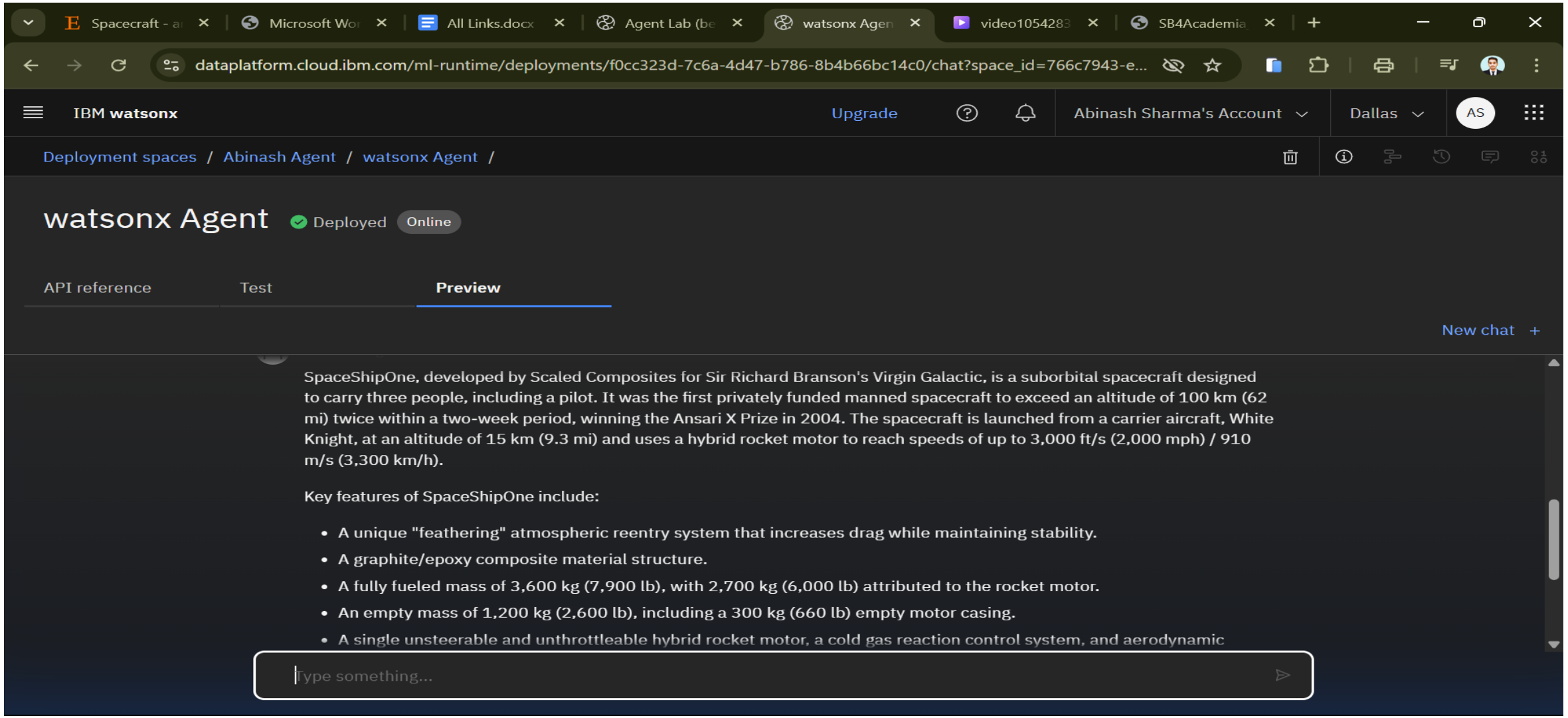



edunet foundation

# RESULT



# RESULT



# CONCLUSION

The Research Agent effectively automates the process of literature search, summarization, and reference management, significantly reducing the time researchers spend on repetitive tasks.

By leveraging **IBM Cloud Lite services** and **IBM Granite models**, the system ensures accurate and context-aware responses suitable for academic and industrial R&D.
During development, challenges included **data preprocessing for diverse document formats** and **optimizing semantic search accuracy**.

These were addressed by implementing vector embeddings and RAG-based retrieval mechanisms.
This solution demonstrates the potential of AI-driven agents to transform research workflows, improving **efficiency, accuracy, and innovation**. Future enhancements can include multi-language support, integration with more academic databases, and real-time collaboration features for research teams.

# FUTURE SCOPE

The Research Agent can be enhanced and expanded in several ways to improve functionality and scalability:

**Incorporation of Additional Data Sources**

- Integrate more academic databases (IEEE, PubMed, Springer) for broader coverage.
- Enable real-time web crawling for the latest research publications.

**Algorithm Optimization**

- Fine-tune IBM Granite models for domain-specific research areas.
- Improve semantic search accuracy using hybrid retrieval techniques.

**Multi-Language & Cross-Domain Support**

- Extend summarization and search to support multiple languages.
- Adapt the system for cross-disciplinary research queries.

**Scalable Deployment**

- Expand deployment to handle larger datasets and concurrent users.
- Implement **edge computing** for faster response in distributed research environments.

**Integration with Emerging Technologies**

- Combine with advanced ML techniques like reinforcement learning for adaptive responses.
- Incorporate **voice-based research assistants** for hands-free querying.

# REFERENCES


IBM Watsonx AI Documentation – https://www.ibm.com/watsonx

IBM Granite Model Overview – https://www.ibm.com/granite

LangChain & LangGraph Documentation – https://python.langchain.com

Lewis, P. et al. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.* NeurIPS.

Devlin, J. et al. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* NAACL-HLT.

# IBM CERTIFICATIONS


edunet foundation

# IBM CERTIFICATIONS


edunet foundation

# IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate



This certificate is presented to

Abinash Sharma

for the completion of

**Lab: Retrieval Augmented Generation with LangChain**

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

**Completion date:** 24 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

# THANK YOU

edunet foundation