
CAPSTONE PROJECT

AI-POWERED RESEARCH ASSISTANT USING IBM WATSONX AND LANGGRAPH

Presented By:

1. Allen John Darin J – Mahendra Engineering College

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

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.

PROPOSED SOLUTION

- The proposed system aims to assist users in automatically conducting literature reviews, discovering relevant research sources, and generating initial content drafts using large language models (LLMs). By leveraging IBM watsonx's foundation models along with LangGraph and web search tools, the agent mimics the cognitive steps a researcher would follow. The solution will consist of the following components:
- **Data Collection:**
 - Integrate real-time access to external data sources using **Google Search**, **Wikipedia Search**, and **Webcrawler** tools to fetch relevant articles.
 - Dynamically collect and analyze content from open-access academic websites, online journals, and science forums.
- **Data Preprocessing:**
 - Extract key text from crawled webpages and filter irrelevant or low-quality information.
 - Structure retrieved information by identifying main concepts, supporting arguments, and source metadata.
 - Maintain context memory to avoid repetition and hallucinations during multi-turn reasoning.
- **Machine Reasoning and Prompt Engineering:**
 - Use **ReAct-style prompting** for step-by-step reasoning and tool invocation.
 - Employ **LangGraph** to create a guided decision path where the agent chooses whether to search, summarize, or draft.
 - Configure **LLaMA 3 70B-Instruct** as the core model for generation, summary, and reasoning.
- **Deployment:**
 - Host the agent on **IBM watsonx.ai Studio** with an intuitive chat interface for user interaction.
 - Allow users to input research queries and receive multi-step responses including source summaries and research directions.
 - Remove unnecessary tools (e.g., Python Interpreter, Document Search) to keep the UI clean and focused.

PROPOSED SOLUTION

Evaluation:

- Evaluate system performance by comparing AI-suggested content with human-curated research outputs.
- Use metrics like **source relevance**, **content originality**, and **coherence of generated text** to assess quality.
- Incorporate user feedback for iterative improvement of prompt instructions and tool order.

Result:

- A fully functional AI research assistant capable of autonomously searching, summarizing, and drafting research content.
- Significantly reduces the time and effort needed for academic or industrial literature reviews.
- Enables non-technical users to access complex research assistance through natural language interaction.

SYSTEM APPROACH

System Requirements

- **Hardware:** Cloud or local system with min. 8-core CPU, 32GB RAM
- **Software:** Python 3.10+, IBM watsonx.ai Studio, Web browser
- **APIs:** Google Search, Wikipedia, Arxiv (optional)

Libraries Used

- `langchain` – agent workflows
- `langgraph` – reasoning graph
- `llama-index` – RAG and structured querying
- `transformers` – LLM access (LLaMA 3)
- `requests`, `bs4` – external search and parsing

Integration Strategy

- Uses **ReAct paradigm** with **LangGraph agents**
- Tools invoked dynamically (search, summarize, draft)
- **LLaMA 3 (watsonx)** generates final responses
- Output: Structured, referenced research summaries

ALGORITHM & DEPLOYMENT

- **Algorithm and Deployment**
- **Algorithm Selection**
Uses the **ReAct (Reasoning + Acting)** framework with **LangGraph** and **LLaMA 3 (70B-Instruct)** for dynamic decision-making and natural language generation. This architecture enables multi-step reasoning, tool use, and autonomous response synthesis — ideal for research workflows.
- **Data Input**
- User research queries in natural language
- External web data from tools like **Google Search**, **Wikipedia**, and **Webcrawler**
- Retrieved page content serves as input context for the LLM
- **Training Process**
- Foundation model **pre-trained on scientific and general knowledge (LLaMA 3)**
- No custom fine-tuning; instead, uses **prompt engineering** and **LangGraph state transitions** to simulate reasoning and adapt to user intent
- **Prediction Process**
- Agent determines required actions: search, summarize, or draft
- Selected tool fetches content
- LLM processes and generates structured output (summary, hypotheses, draft section) based on the latest retrieved context

RESULT

Tools

Add a tool

Create custom tool

Added tools (3)

Google search

Retrieve information from the Internet with the Google search engine.



Wikipedia search

Retrieve information from Wikipedia articles.



Webcrawler

Retrieve information from a website.



Integrated tools include Google Search, Wikipedia, and a Webcrawler to retrieve up-to-date and comprehensive information.

Your capabilities include:

- Understanding research queries in natural language.
- Searching online resources using Google and Wikipedia.
- Extracting and summarizing relevant content from web pages.
- Suggesting research directions or basic paper sections.

The agent is configured with natural language understanding and web-search capabilities, enabling it to support research queries effectively

RESULT

The screenshot displays the IBM watsonx Agent web interface. The browser address bar shows the URL: `dataplatform.cloud.ibm.com/wx/agents/f7b5dd44-5e2c-45a1-b31a-1fe4c4785c2?project_id=48d4ee3f-2a75-445c-8c9e-a3b85d0a82a7&context=wx`. The interface is divided into two main sections: 'Build' on the left and 'Agent preview' on the right.

Build Section:

- Model:** llama-3-3-70b-instruct
- Framework:** LangGraph
- Architecture:** ReAct
- Instructions:** You are a research assistant AI agent designed to help with academic and scientific research. Your capabilities include:
 - Understanding research queries in natural language.
 - Searching online resources using Google and Wikipedia.
- Knowledge:** (Empty section)
- Tools:** (Empty section with buttons 'Add a tool' and 'Create custom tool')

Agent preview Section:

User: 08:19 PM
Find the latest research articles on "quantum machine learning applications in drug discovery."

watsonx Agent: 08:19 PM
Based on the search results, here are some potential sources for the latest research articles on "quantum machine learning applications in drug discovery":

1. "Quantum Machine Learning in Drug Discovery: Applications in Academia and Pharmaceutical" (arXiv paper 2409.15645)
2. "Quantum Machine Learning Algorithms for Drug Discovery" (PMC article)
3. "Quantum Machine Learning in Drug Discovery: Applications in Academia and Pharmaceutical" (ACS article)
4. "Unlocking the Potential of Quantum Machine Learning to Advance Drug Discovery" (MDPI article)
5. "Quantum Machine Learning Algorithms for Drug Discovery" (ACS article)
6. "Hybrid quantum-classical machine learning for generative chemistry" (Nature article)
7. "Quantum Machine Learning in Drug Discovery: Current State and Future Directions" (ACM article)
8. "Synergizing quantum techniques with machine learning for drug discovery" (Nature article)

These articles and papers discuss various aspects of quantum machine learning in drug discovery.

Input field: |type something... |

Demonstration of the assistant responding to a research query by listing relevant academic sources on 'quantum machine learning in drug discovery.'

CONCLUSION

- The Research Agent leverages IBM watsonx and Granite LLMs to automate literature review, summarization, and citation tasks. It improves research efficiency by reducing manual effort and enabling rapid knowledge synthesis. While accuracy and contextual understanding can still be enhanced, the system demonstrates significant potential to transform both academic and industrial R&D processes.

FUTURE SCOPE

- **Broader Data Integration:** Add more sources like patent databases, preprints, and institutional repositories.
- **Enhanced Algorithms:** Incorporate advanced models like GPT-4o or retrieval-augmented generation (RAG) for deeper insight.
- **Cross-Domain Expansion:** Adapt system for multiple disciplines beyond technical research (e.g., medical, legal).
- **Real-Time Updates:** Enable dynamic ingestion of newly published work.
- **Edge & Cloud Deployment:** Explore edge computing for faster, decentralized processing in collaborative environments.

REFERENCES

- Zhang et al., “A Survey on AI for Scientific Discovery,” *Nature Machine Intelligence*, 2021.
- IBM watsonx Documentation – <https://www.ibm.com/docs/en/watsonx>
- Liu et al., “Natural Language Processing for Literature Review Automation,” *ACM Computing Surveys*, 2022.
- BERT: Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *arXiv:1810.04805*, 2019.
- Open Research Corpus – Semantic Scholar (AllenAI)

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Allen John Darin J

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence



Issued on: Jul 18, 2025
Issued by: IBM SkillsBuild

Verify: <https://www.credly.com/badges/a2558c14-8a7f-4ed1-a90c-a85f7bc98f19>



IBM CERTIFICATIONS



IBM CERTIFICATIONS





THANK YOU