
CAPSTONE PROJECT

INTELLEXA THE INTELLIGENT RESEARCH ASSISTANT

Presented By:

- 1. Student Name: Vadla Yogeshwar**
- 2. AICTE Student ID: STU6596b38a25dfe1704375178**
- 3. College Name-Department: Malla Reddy College of Engineering and Technology – CSE – Data Science**

OUTLINE

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

PROBLEM STATEMENT

- Academic and scientific research involves various complex and repetitive tasks such as literature search, paper summarization, citation management, and reference organization. These processes are time-consuming and often reduce the efficiency and productivity of researchers.
- There is a growing need for an AI-based solution that can assist researchers by automating these tasks. A Research Agent is designed to serve this purpose by leveraging natural language processing and foundation models. It can autonomously search for literature, understand research questions, summarize academic papers, suggest hypotheses, generate content for reports, and manage citations.
- The objective is to develop an intelligent Research Agent that enhances the research process by improving accuracy, saving time, and supporting innovation across academic and industrial research environments.

PROPOSED SOLUTION

To address the challenges faced in academic and scientific research workflows, the proposed solution is to develop an AI-powered **Research Agent** named *Intellexa*. This intelligent assistant utilizes IBM Watsonx foundation models and natural language processing capabilities to automate and enhance various research-related tasks.

- The system is designed to:
- Accept user input in the form of research queries, abstracts, or topics.
- Use a foundation model (such as IBM Granite) to process and understand natural language input.
- Generate precise summaries of academic papers.
- Suggest relevant research hypotheses based on provided abstracts or topics.
- Organize references and generate citations in standard formats (e.g., APA, MLA).
- Respond with structured and step-by-step explanations to complex research questions.
- The Research Agent is built and tested using IBM Cloud services and deployed in a Jupyter notebook environment, allowing future integration into broader academic platforms or web-based applications. This solution aims to significantly reduce manual effort, increase research accuracy, and streamline academic workflows.

SYSTEM APPROACH

The system approach outlines the overall strategy and methodology followed to develop and implement the Intellexa Research Agent. The goal is to leverage large language models (LLMs) through IBM Watsonx to automate various academic research tasks such as summarization, hypothesis generation, and citation formatting.

System requirements

- **Hardware Requirements:**
 - Internet-enabled device (PC/Laptop)
 - Minimum 8 GB RAM
 - Web browser (Chrome, Firefox, Edge)
- **Software/Platform Requirements:**
 - IBM Cloud account with access to Watsonx.ai
 - IBM Project environment (to store agents and notebooks)
 - Jupyter Notebook environment (local or cloud-based)

■ Library required to build the model

The following tools and libraries are used in building the agent and notebook:

IBM Cloud Services:

- Watsonx.ai (Prompt Lab)
- IBM Foundation Models (Granite)

Python Libraries:

- ibm-watson-machine-learning: For authentication and service access
- langchain_ibm: To integrate LangChain with IBM models
- langgraph: For building tool-based agent workflows
- requests: To make API calls if needed
- dotenv: To manage environment variables securely
- IPython.display: For better display in Jupyter output

■ Methodology

- Define agent behavior using system instructions and prompt templates
- Select appropriate LLM (Granite 3 – 2B)
- Use Prompt Lab to design and test responses
- Save the agent as a notebook and download it for integration

ALGORITHM & DEPLOYMENT

Algorithms

- The core of the Intellexa Research Agent is built on a **Large Language Model (LLM)** rather than a conventional machine learning algorithm. The selected model (such as Granite 3 – 2B) is a pre-trained foundation model capable of performing a variety of natural language understanding and generation tasks.

Key characteristics of the algorithmic foundation:

- **Transformer-based Architecture**
The underlying model is built using an optimized transformer architecture, which processes input in parallel and captures contextual relationships between words in long texts.
- **Instruction-Tuned Behavior**
The model is instruction-following, meaning it has been fine-tuned on datasets where it learns to follow user commands or prompts effectively.
- **Zero-shot and Few-shot Learning**
The system does not require task-specific training. It generalizes across a wide range of research-oriented tasks with minimal or no example data.

Deployment Strategy

The deployment of the Intellexa agent is handled using IBM Cloud's **Watsonx.ai** and **Jupyter Notebook** environments.

Steps for Deployment:

- **Model Selection and Configuration**

The LLM (Granite 3 – 2B) is selected from the Watsonx Prompt Lab and configured with system instructions that define agent behavior.

- **Agent Design in Prompt Lab**

The research agent is built and tested in Prompt Lab using prompt engineering techniques.

- **Notebook Export**

The agent is saved as a **Standard Notebook (.ipynb)** containing the setup and logic.

- **Local Execution**

The notebook can be downloaded and executed locally in any Jupyter environment to interact with the agent using real-time prompts.

- **Future Scope: API-based Deployment**

The notebook includes code stubs for integrating the model via REST API, enabling future deployment as:

- A web application
- A chatbot interface
- A research assistant plugin in academic platforms

RESULT

The Intellexa Research Agent was evaluated based on its ability to handle various research-related tasks through natural language interactions. Since the agent is built on top of a Large Language Model (LLM), its performance is measured in terms of:

1. Effectiveness of Responses

- The agent successfully summarized academic abstracts and papers with clarity and relevance.
- It generated meaningful research hypotheses from minimal input.
- It provided correct citation formatting in APA style without error.
- The system responded promptly and accurately to complex queries.


2. Consistency

- Responses remained consistent across multiple queries with similar intent.
- The system handled follow-up prompts well, maintaining conversational context.


3. GitHub Link: https://github.com/Yogeshwarypl/Intellexa_The-Intelligent_Research_Assistant.git

effectiveness of the investigation process with data privacy preservation.

► How did I get this answer?

 You: 08:57 PM

What are some recent studies related to my topic?

 Intellexa: The Intelligent Research Assistant 08:57 PM

Based on the search results, here are some recent studies related to digital forensic investigations and data privacy:

1. "Privacy impact assessment in large-scale digital forensic investigations" - This study describes how to conduct a privacy impact assessment in large-scale digital forensic investigations.
2. "Investigating Wearable Fitness Applications: Data Privacy and Digital Forensics" - This study investigates the data privacy and digital forensics of wearable fitness applications.
3. "What Is Digital Forensics? A Closer Examination of the Field" - This article provides an overview of the field of digital forensics and its current applications in criminal and corporate investigations.
4. "M.S. Information Security and Digital Forensics Online (ISDF)" - This is a master's program in information security and digital forensics that covers topics such as data privacy and digital forensics.
5. "Chapter 2 DATA PRIVACY PERCEPTIONS ABOUT DIGITAL FORENSIC INVESTIGATIONS" - This study collects the ground truth about the principal data privacy issues related to digital forensic investigations.
6. "Privacy-Preservation and Containment in IoT Forensics Investigations: A Comparative Study" - This study compares different approaches to privacy preservation and containment in IoT forensics investigations.
7. "The Role of AI in Forensics" - This article discusses the role of artificial intelligence in digital forensics and its potential to enhance data privacy and security.
8. "Protecting Digital Data Privacy in Computer Forensic Examination" - This study presents a cryptographic model designed to protect digital data privacy in computer forensic examinations.
9. "Cyber Center for Security and Analytics" - This is a research center that conducts high-impact, applied cyber security and data science research, development, operations, and training.
10. "Data privacy preservation in digital forensics investigation" - This study analyzes the perceptions about data privacy preservation and issues related to digital forensics investigations.

These studies demonstrate the ongoing research and development in the field of digital forensics and data privacy, and highlight the importance of protecting data privacy in digital forensic investigations.

Type something...



Agent preview

try to summarize it for you.

```
{"name": "WebCrawler", "parameters": {"url": "valid research paper URL"}}
```

► How did I get this answer?



You: 08:56 PM

<https://commons.erau.edu/jdfs/vol14/iss2/3> Can you summarize this research paper for me?



Intellexa: The Intelligent Research Assistant 08:56 PM

The research paper "DF 2.0: An Automated, Privacy Preserving, and Efficient Digital Forensic Framework That Leverages Machine Learning for Evidence Prediction and Privacy Evaluation" discusses the challenges faced by digital forensic investigators in handling large amounts of data and maintaining data privacy. The authors propose a framework called DF 2.0, which uses machine learning to predict evidence and evaluate privacy, while also automating the digital forensic analysis process. The framework aims to improve the efficiency of digital forensic investigations while preserving data privacy.

The paper highlights the importance of data privacy in digital forensic investigations and discusses the current state of digital forensic research. It also provides an overview of the DF 2.0 framework, including its architecture and components. The authors discuss the benefits of using machine learning in digital forensic investigations, including improved accuracy and efficiency.

The paper also discusses the challenges faced by digital forensic investigators, including the increasing volume of digital data, the complexity of digital devices, and the need to maintain data privacy. The authors argue that the current state of digital forensic research is focused on either the effectiveness of the investigation process or data privacy preservation, but not both. They propose that the DF 2.0 framework can address these challenges by providing a comprehensive approach that preserves data privacy while improving the efficiency of digital forensic investigations.

Overall, the paper provides a detailed overview of the DF 2.0 framework and its potential benefits for digital forensic investigations. It highlights the importance of data privacy and the need for a comprehensive approach that balances the effectiveness of the investigation process with data privacy preservation.

► How did I get this answer?



Type something...



CONCLUSION

The development of *Intellexa – The Intelligent Research Assistant* demonstrates the potential of leveraging large language models (LLMs) to automate and enhance various academic research tasks. By integrating IBM Watsonx foundation models within a prompt-driven agent framework, the system effectively performs literature summarization, hypothesis generation, and citation formatting without the need for traditional machine learning model training.

The results show that the agent can understand complex research prompts and provide contextually relevant and structured responses, improving research productivity and reducing manual effort.

■ Challenges Encountered:

- Limited customization of LLM behavior beyond prompt engineering
- Constraints in downloading deployment-ready notebooks in some environments
- Handling model limitations when responses require highly domain-specific knowledge

■ Potential Improvements:

- Integration with academic databases or research APIs (e.g., IEEE, PubMed)
- Addition of PDF parsing capabilities for full paper summarization
- Enhanced prompt templates for domain-specific fields (e.g., medical, legal)
- This project demonstrates the growing role of AI-powered agents in academic environments and provides a foundation for building more advanced, intelligent research tools in the future.

FUTURE SCOPE

The Intellexa Research Agent provides a foundational step toward automating academic research tasks using large language models. However, there is considerable potential for enhancing and expanding the system in the future. Key areas for development include:

1. Integration with External Data Sources

- Connect the agent to academic databases such as IEEE Xplore, SpringerLink, PubMed, or Google Scholar to enable direct literature search and summarization from live sources.
- Allow ingestion of PDF research papers for end-to-end document understanding and processing.

2. Domain-Specific Customization

- Fine-tune or adapt prompts for domain-specific use cases (e.g., medical research, engineering, legal studies).
- Introduce specialized toolchains for each domain to improve relevance and accuracy.

3. Improved Interaction and Context Handling

- Implement conversational memory to allow follow-up questions and multi-turn dialogue.
- Add dynamic context tracking for long-form discussions or continuous research sessions.

4. Deployment and Scalability

- Deploy the agent as a full-stack web application with user login, history tracking, and file upload.
- Use containerization tools (e.g., Docker) for scalable deployment on cloud or on-premise servers.

5. Integration of Advanced Technologies

- Incorporate **edge computing** for offline or low-latency deployment in institutional environments.
- Explore **multi-modal models** that can analyze both text and visual content (e.g., research figures, charts).
- Utilize **retrieval-augmented generation (RAG)** techniques for combining knowledge bases with model outputs.
- The long-term vision is to transform Intellexa into a comprehensive AI research assistant that can support academics and professionals across multiple disciplines, institutions, and geographical regions.

REFERENCES

- IBM watsonx.ai Documentation. *IBM Knowledge Center*. <https://www.ibm.com/cloud/watsonx-ai>
- Touvron, H., Lavril, T., Izacard, G., et al. (2023).
LLaMA 2: Open Foundation and Fine-Tuned Chat Models.
Meta AI Research.
<https://arxiv.org/abs/2307.09288>
- LangChain Documentation.
<https://docs.langchain.com/>
- LangGraph for Agent Workflows.
<https://github.com/langchain-ai/langgraph>
- OpenAI. (2023).
GPT-4 Technical Report.
<https://cdn.openai.com/papers/gpt-4.pdf>
- IBM Cloud -Getting Started with Jupyter Notebooks.
<https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/notebooks-parent.html>
- IBM Granite Models Overview.
<https://www.ibm.com/blog/what-are-granite-models/>

IBM CERTIFICATIONS



IBM CERTIFICATIONS

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to

Yogeshwar Vadla

for the completion of

Journey to Cloud: Envisioning Your Solution

(MDL-447)

According to the Moodle system of record

Completion date: 31 Jul 2025 (GMT)

Learning hours: 5 hrs

IBM CERTIFICATIONS





THANK YOU