Functional Search Engine for Agriculture

Leveraging LLM and Azure Services

Date: December 20, 2024

Overview

This project aims to develop a robust search engine powered by a Large Language Model (LLM).
The application integrates various Azure services to enhance user interaction and provide accurate search results based on user queries. Currently, the implementation focuses on leveraging LLM capabilities to retrieve and generate relevant information efficiently.

Developed by:
Adriano G. Back End Development & Integration
Danillo D. Back End Development & Integration
Rodica V. Front End Development & Project Planning and Presentation
Demerson P. Front End & Back End Integration and Research planning

1 Key Features

Functional Search Engine: The application operates as a search engine, utilizing LLM to process user queries and return relevant results from a vast database of information.

Data Source: The project utilizes a data source from Brazil, allowing users to ask questions exclusively in Portuguese. Responses from OpenAI can be provided in both Portuguese and English, ensuring bilingual support.

Integration of Azure Services: The project employs a variety of Azure services to enhance functionality, including:

Azure OpenAI: For generating human-like responses and understanding user queries.

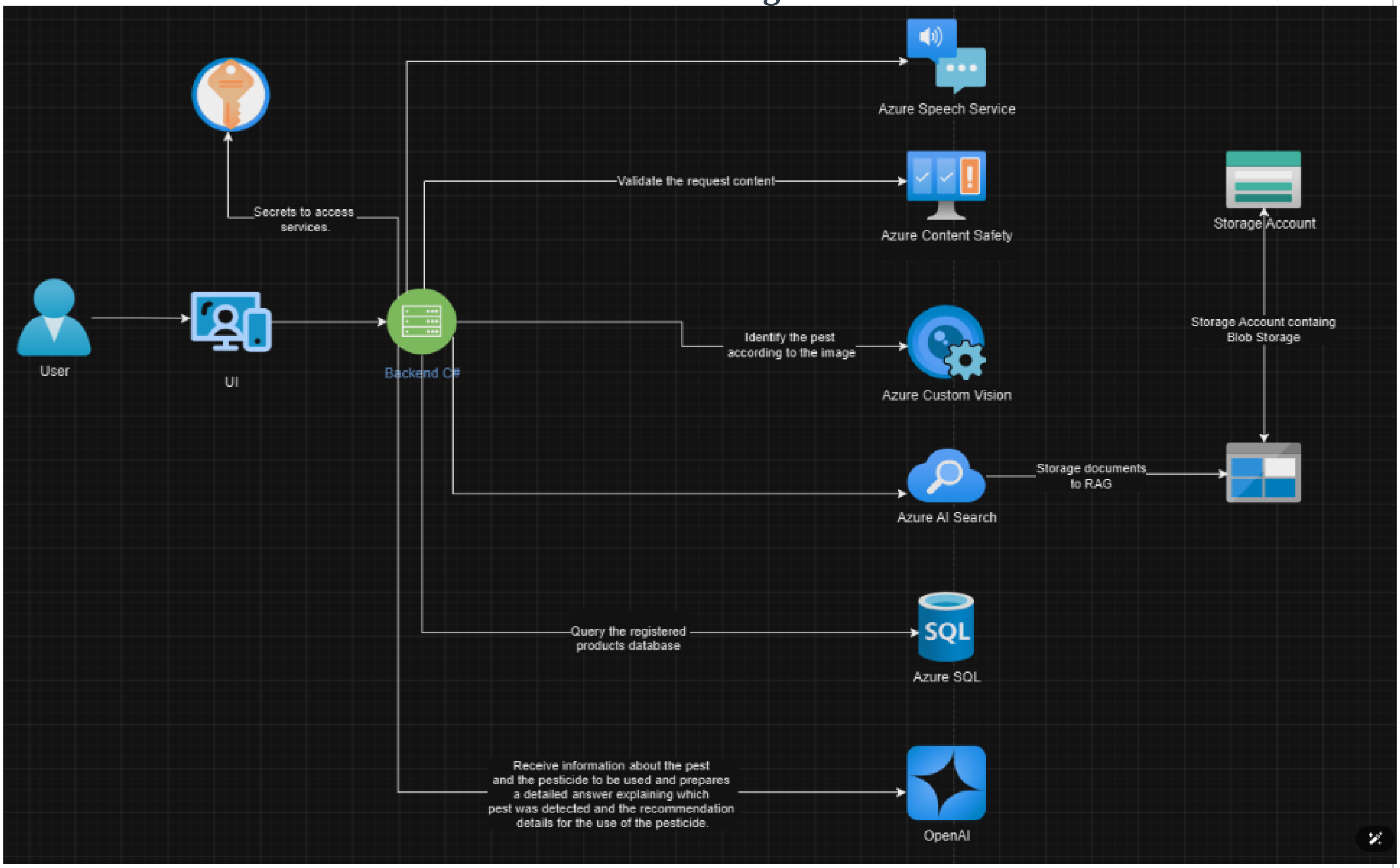
Azure Cosmos DB: To store and manage large datasets efficiently.

Azure Content Moderator: To ensure that the content returned adheres to community standards and guidelines.

Azure Speech to Text: To allow voice-based queries, enhancing accessibility.

Azure Custom Vision: For image recognition capabilities, enabling users to search using images.

Workflow Diagram



2 Language Extensibility

Future Capabilities: The project can be extended to support additional languages, including English and others, provided that appropriate data is stored in the database. This adaptability allows for broader application across different countries and linguistic demographics.

CI/CD and Azure DevOps

Implementation Status: Due to time constraints, we did not have the opportunity to implement Continuous Integration and Continuous Deployment (CI/CD) practices or utilize Azure DevOps for this project.

Understanding of Software Lifecycle: Despite the lack of CI/CD implementation, the team possesses a solid understanding of the entire software development lifecycle. We have adhered to best practices throughout the development process.

Version Control: We utilized GitHub repositories for version control, ensuring effective collaboration and tracking of changes.

3 Responsible AI

We emphasize responsible AI practices by implementing Azure Content Moderator on user inputs and providing disclaimers on outputs. This ensures ethical standards are maintained and potential risks associated with AI usage are mitigated.

4 Innovation and Future Development

While the current implementation serves as a functional search engine, there is significant potential for further innovation. Future enhancements could include:

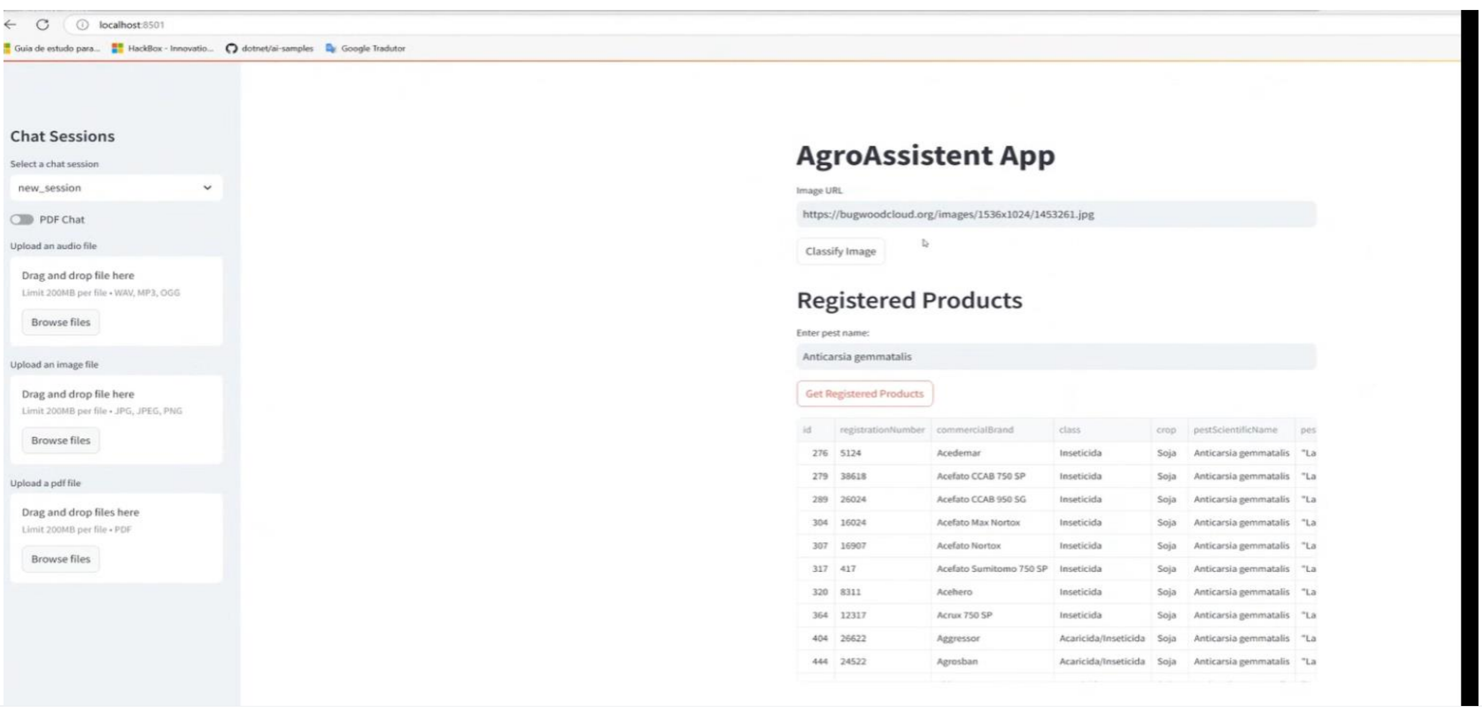
Advanced Query Processing: Implementing more sophisticated algorithms for understanding user intent and context.

Integration of Additional Data Sources: Expanding the search capabilities by integrating external databases and APIs.

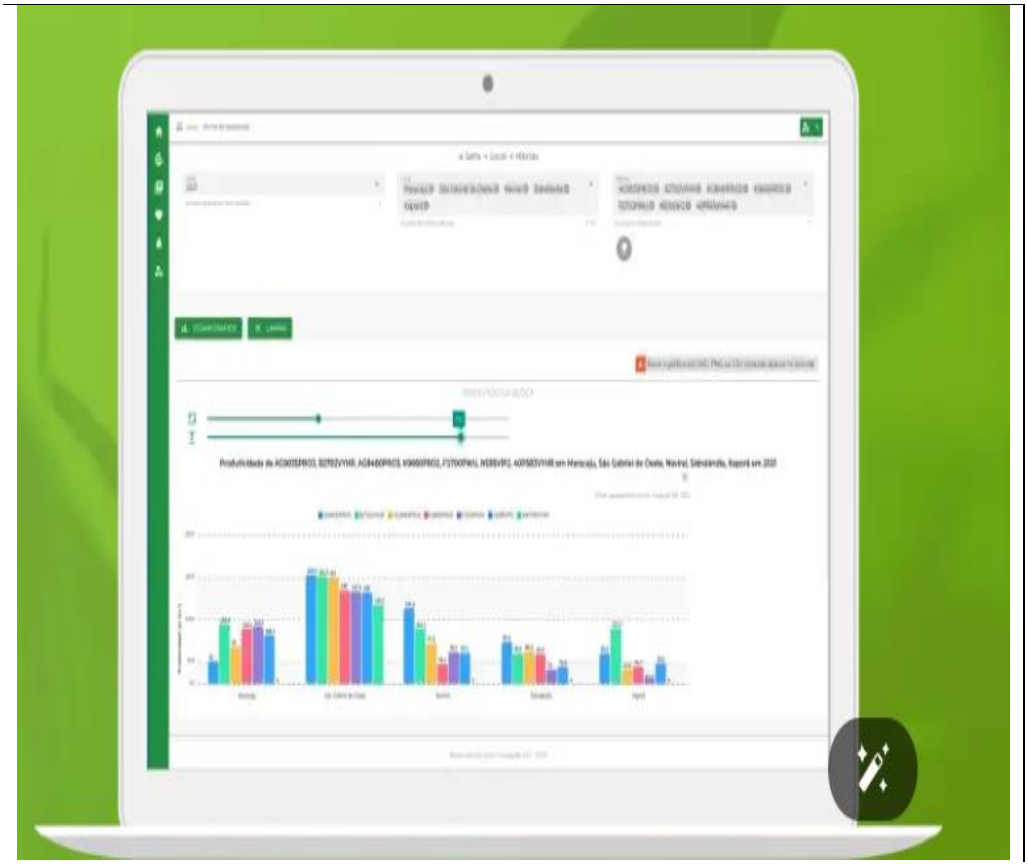
Report Generation: Developing features that allow users to generate reports based on search results, particularly in fields like agriculture, where data-driven insights are crucial.

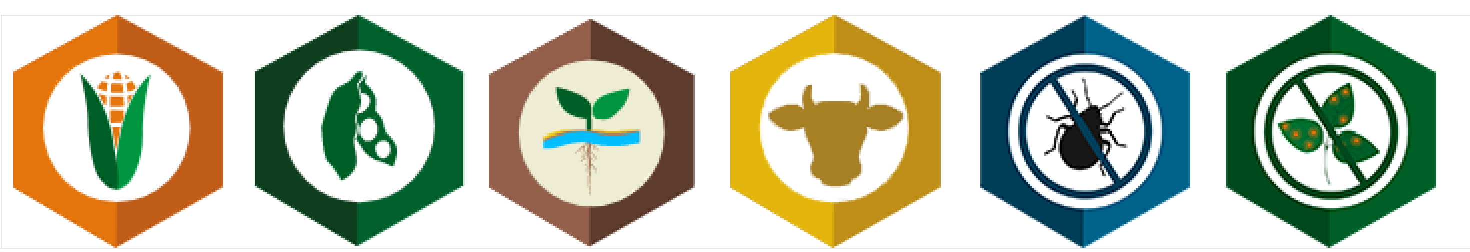
Validation of the use of technology in compliance with laws.

The project was developed based on a **Strategic Agronomic Planning** program, which is informed by the research conducted by **Foundation researchers**. This program focuses on selecting the most effective inputs for optimizing crop production.

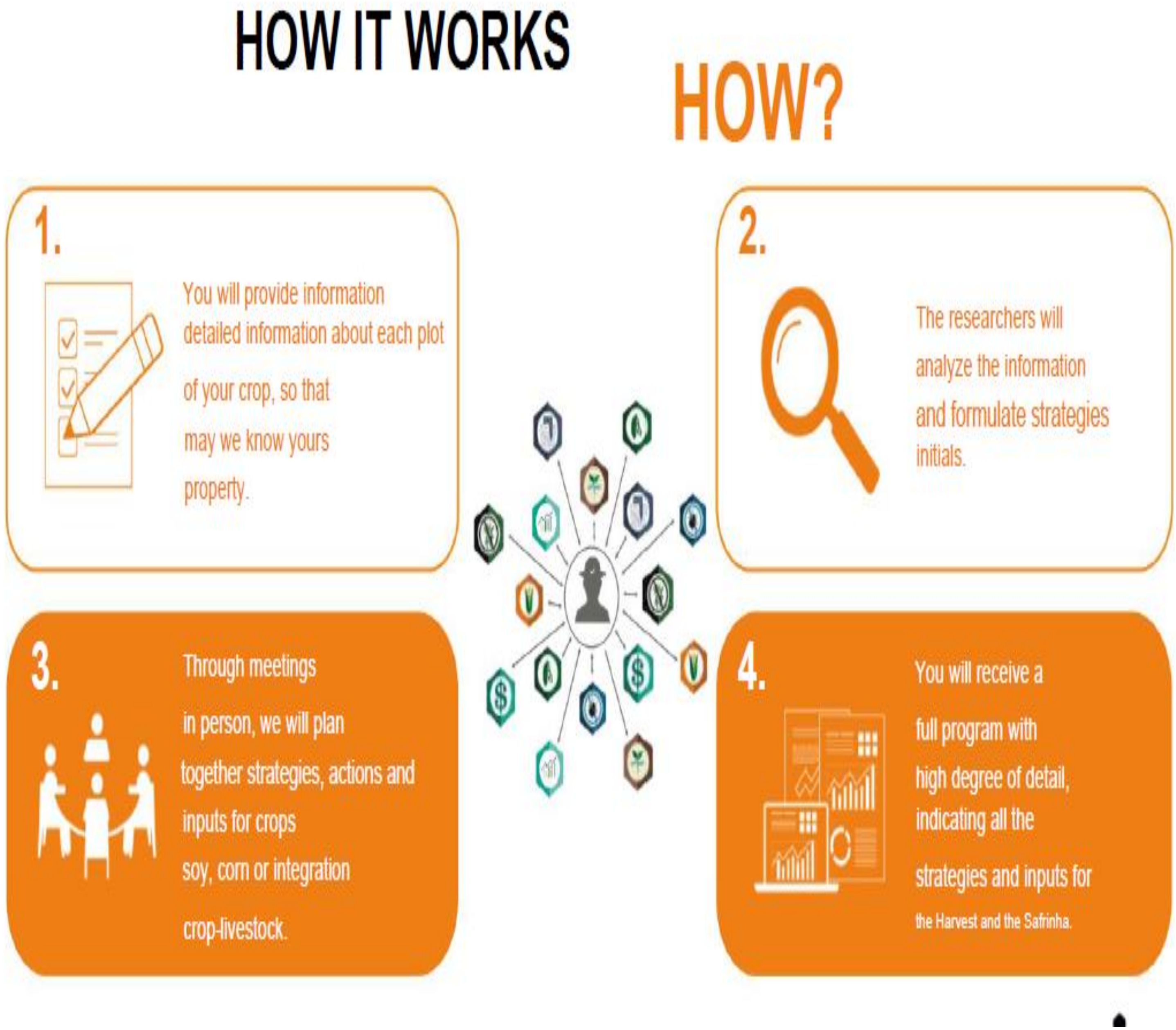


BECOME A MEMBER
BOOST YOUR PRODUCTIVITY AND DRIVE MORE VALUE FOR YOUR BUSINESS

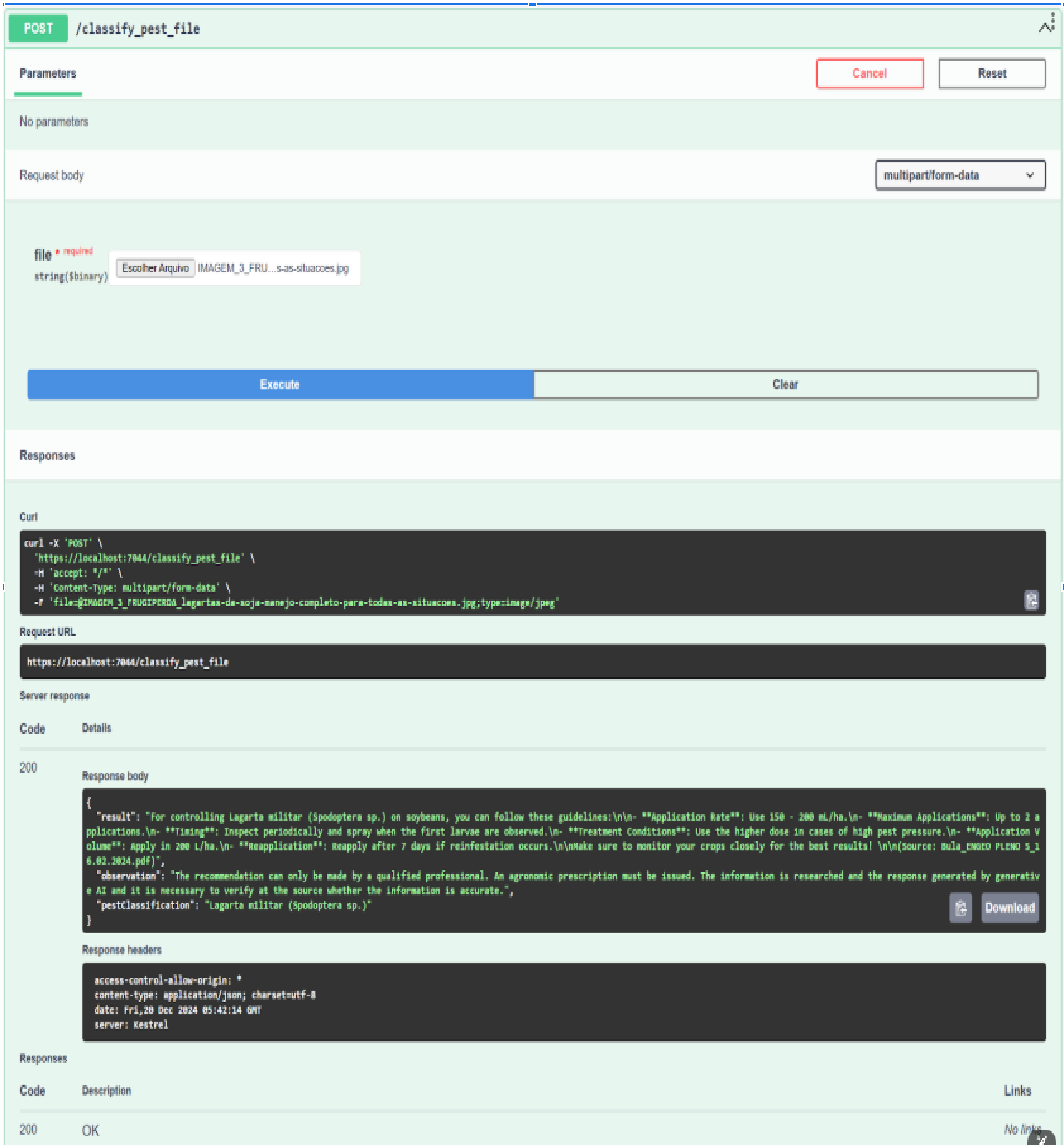




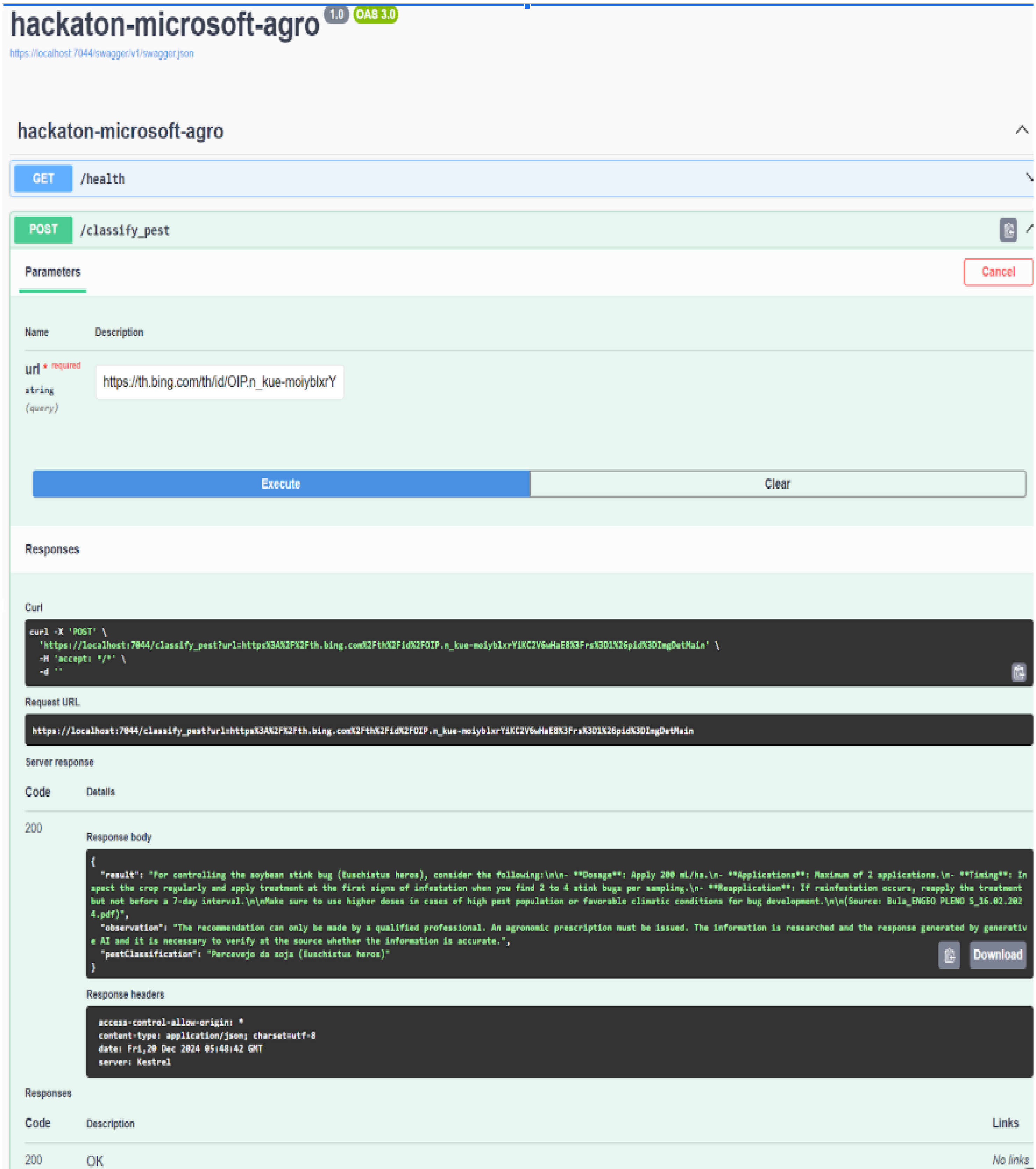
PLANT SCIENCE CORN AND SORGHUM SOYBEAN PHYTOTECHNICS SOIL MANAGEMENT AND FERTILITY INTEGRATED PRODUCTION SYSTEMS HERBOLOGIA E ENTOMOLOGIA PHYTOPATHOLOGY AND NEMATOLOGY

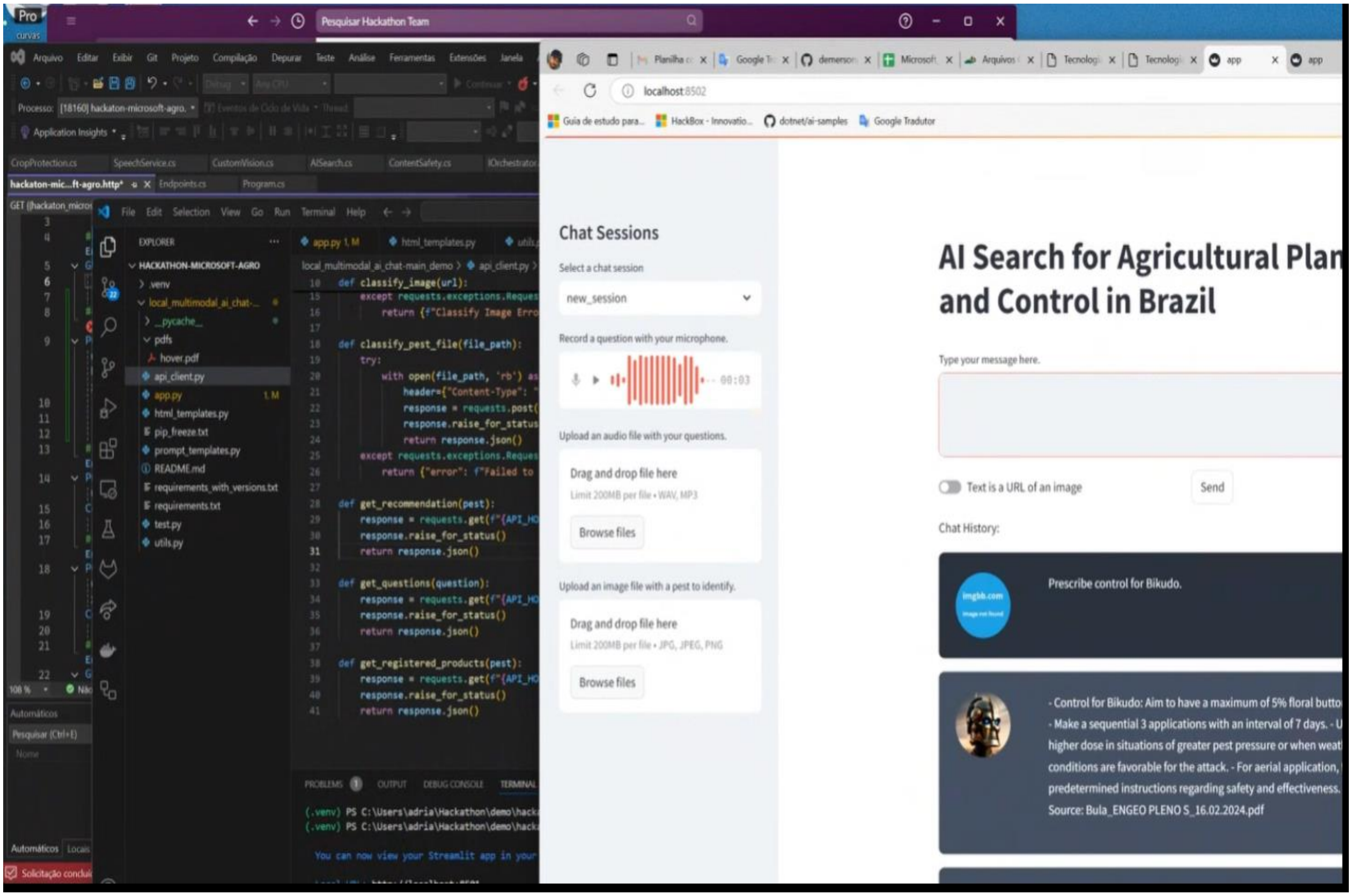


Using **Swagger**, you can test **GET** commands to retrieve data and **POST** commands to submit data, while easily viewing detailed responses, including status codes and response bodies, for each **API endpoint**.



POST
/classify_pest_file
Parameters
Cancel
Reset
No parameters
Request body
multipart/form-data
file * required
string($binary)
Escolher Arquivo
IMAGEM_3_FRU...s-as-situacoes.jpg
Execute
Clear
Responses
Curl
curl -X 'POST' \
'https://localhost:7044/classify_pest_file' \
-H 'accept: */*' \
-H 'Content-Type: multipart/form-data' \
-F 'file=@IMAGEM_3_FRUGIPERDA_lagartas-da-soja-manejo-completo-para-todas-as-situacoes.jpg;type=image/jpeg'
Request URL
https://localhost:7044/classify_pest_file
Server response
Code
Details
200
Response body
{
"result": "For controlling Lagarta militar (Spodoptera sp.) on soybeans, you can follow these guidelines:\n\n- **Application Rate**: Use 150 - 200 mL/ha.\n- **Maximum Applications**: Up to 2 a
pplications.\n- **Timing**: Inspect periodically and spray when the first larvae are observed.\n- **Treatment Conditions**: Use the higher dose in cases of high pest pressure.\n- **Application V
olume**: Apply in 200 L/ha.\n- **Reapplication**: Reapply after 7 days if reinfestation occurs.\n\nMake sure to monitor your crops closely for the best results! \n\n(Source: Bula_ENGEO PLENO S_1
6.02.2024.pdf)",
"observation": "The recommendation can only be made by a qualified professional. An agronomic prescription must be issued. The information is researched and the response generated by generativ
e AI and it is necessary to verify at the source whether the information is accurate.",
"pestClassification": "Lagarta militar (Spodoptera sp.)"
}
Download
Response headers
access-control-allow-origin: *
content-type: application/json; charset=utf-8
date: Fri,20 Dec 2024 05:42:14 GMT
server: Kestrel
Responses
Code
Description
Links
200
OK
No links


hackaton-microsoft-agro
1.0
OAS 3.0
https://localhost:7044/swagger/v1/swagger.json
hackaton-microsoft-agro
GET
/health
POST
/classify_pest
Parameters
Cancel
Name
Description
url * required
string
(query)
https://th.bing.com/th/id/OIP.n_kue-moiyblxrY
Execute
Clear
Responses
Curl
curl -X 'POST' \
'https://localhost:7044/classify_pest?url=https%3A%2F%2Fth.bing.com%2Fth%2Fid%2FOIP.n_kue-moiyblxrYiKC2V6wHaE8%3Frs%3D1%26pid%3DImgDetMain' \
-H 'accept: */*' \
-d ''
Request URL
https://localhost:7044/classify_pest?url=https%3A%2F%2Fth.bing.com%2Fth%2Fid%2FOIP.n_kue-moiyblxrYiKC2V6wHaE8%3Frs%3D1%26pid%3DImgDetMain
Server response
Code
Details
200
Response body
{
"result": "For controlling the soybean stink bug (Euschistus heros), consider the following:\n\n- **Dosage**: Apply 200 mL/ha.\n- **Applications**: Maximum of 2 applications.\n- **Timing**: Inspect the crop regularly and apply treatment at the first signs of infestation when you find 2 to 4 stink bugs per sampling.\n- **Reapplication**: If reinfestation occurs, reapply the treatment but not before a 7-day interval.\n\nMake sure to use higher doses in cases of high pest population or favorable climatic conditions for bug development.\n\n(Source: Bula_ENGEO PLENO S_16.02.2024.pdf)",
"observation": "The recommendation can only be made by a qualified professional. An agronomic prescription must be issued. The information is researched and the response generated by generative AI and it is necessary to verify at the source whether the information is accurate.",
"pestClassification": "Percevejo da soja (Euschistus heros)"
}
Download
Response headers
access-control-allow-origin: *
content-type: application/json; charset=utf-8
date: Fri,20 Dec 2024 05:48:42 GMT
server: Kestrel
Responses
Code
Description
Links
200
OK
No links



Regions exhibit significant variations in temperature, rainfall, day length, and other environmental factors, all of which influence plant behavior and crop production. Additionally, the timing of planting periods plays a crucial role in determining plant yield.

Given the vast amount of information available, there is a substantial need to compile and organize this data into a comprehensive database.

Generative AI can facilitate this process by enhancing search capabilities. With the abundance of productivity research published globally, AI can comprehend context and deliver more relevant search results, making it easier to access the information needed for effective agronomic planning. or effective agronomic

Tabela 1 - Épocas de semeaduras de cultivares de soja na Região Centro-Sul do estado de Mato Grosso do Sul. Fonte: Adaptado de Pitol (2015).

Cultivar	Época de Plantio - Dia/Mês																			
	Setembro		Outubro						Novembro						Dezembro					
	20 a 25	26 a 30	01 a 05	06 a 10	11 a 15	16 a 20	21 a 25	26 a 30	01 a 05	06 a 10	11 a 15	16 a 20	21 a 25	26 a 30	01 a 05	06 a 10	11 a 15	16 a 20	21 a 25	26 a 30
5D6215-IPRO																				
ADV4317-IPRO																				
AS3610-IPRO																				
AS3680-IPRO																				
AS3730-IPRO																				

Estande: aumenta ← menor → aumenta
Altura: diminui ← maior → diminui

Período preferencial = alta produtividade + competitividade

Período tolerado = média produtividade/risco de estiagem e doença

Período aceitável = boa produtividade

Não recomendado = baixa profundidade/baixo porte/alto riscos com doenças e estiagem

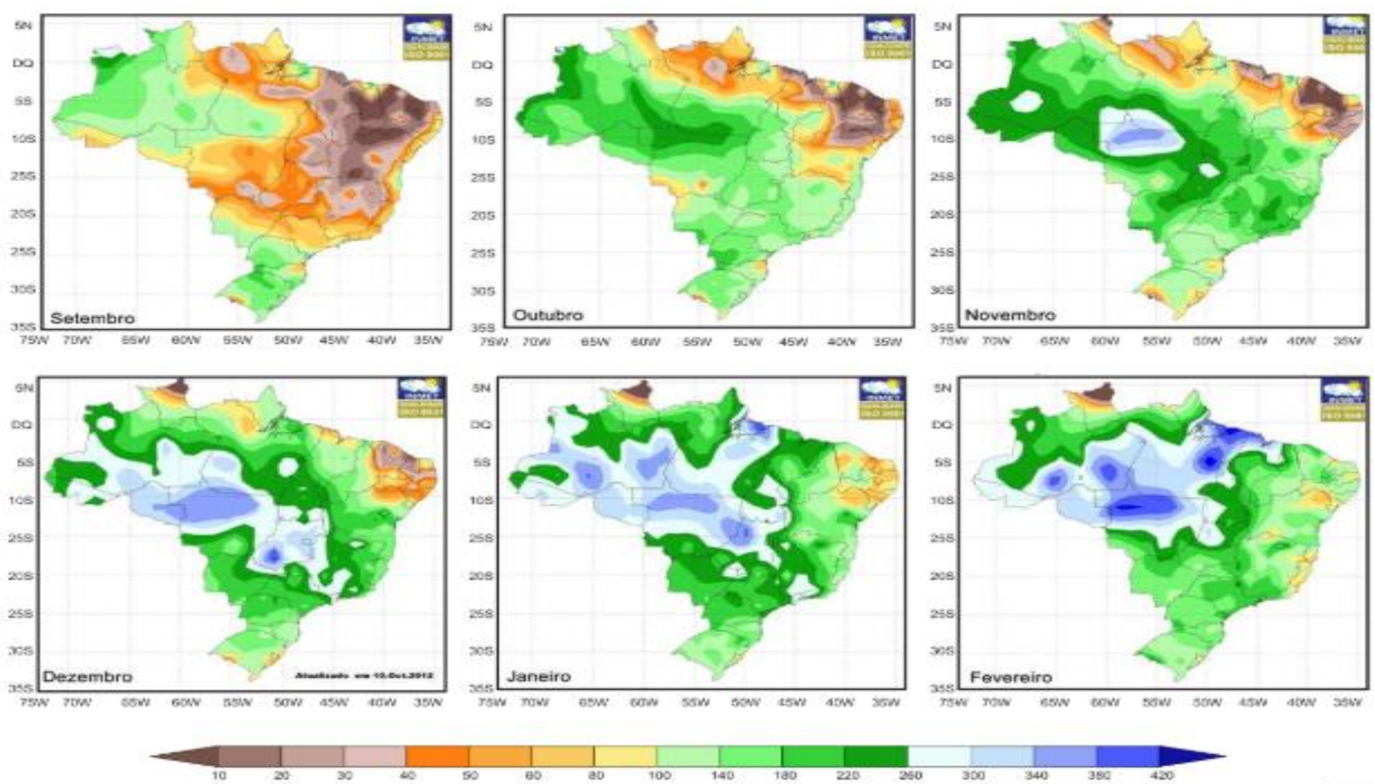


Figura 4 - Precipitação acumulada mensal nos meses de setembro a fevereiro com base nas normas climatológicas do Brasil (1961-1990). Fonte: INMET

Different cultivars, which are plants with varying genetic traits, respond uniquely to specific environmental conditions. Consequently, crop production can differ significantly from one region to another.

The following table illustrates the percentage of weed control achieved with various herbicides, a trend that is also applicable to fungi and insect management. For instance, one herbicide may provide 100% control of a particular pest while only managing to control 15% of another. Conversely, another option might achieve 80% control of both pests. Additionally, varying levels of pest infestation can further complicate these dynamics.

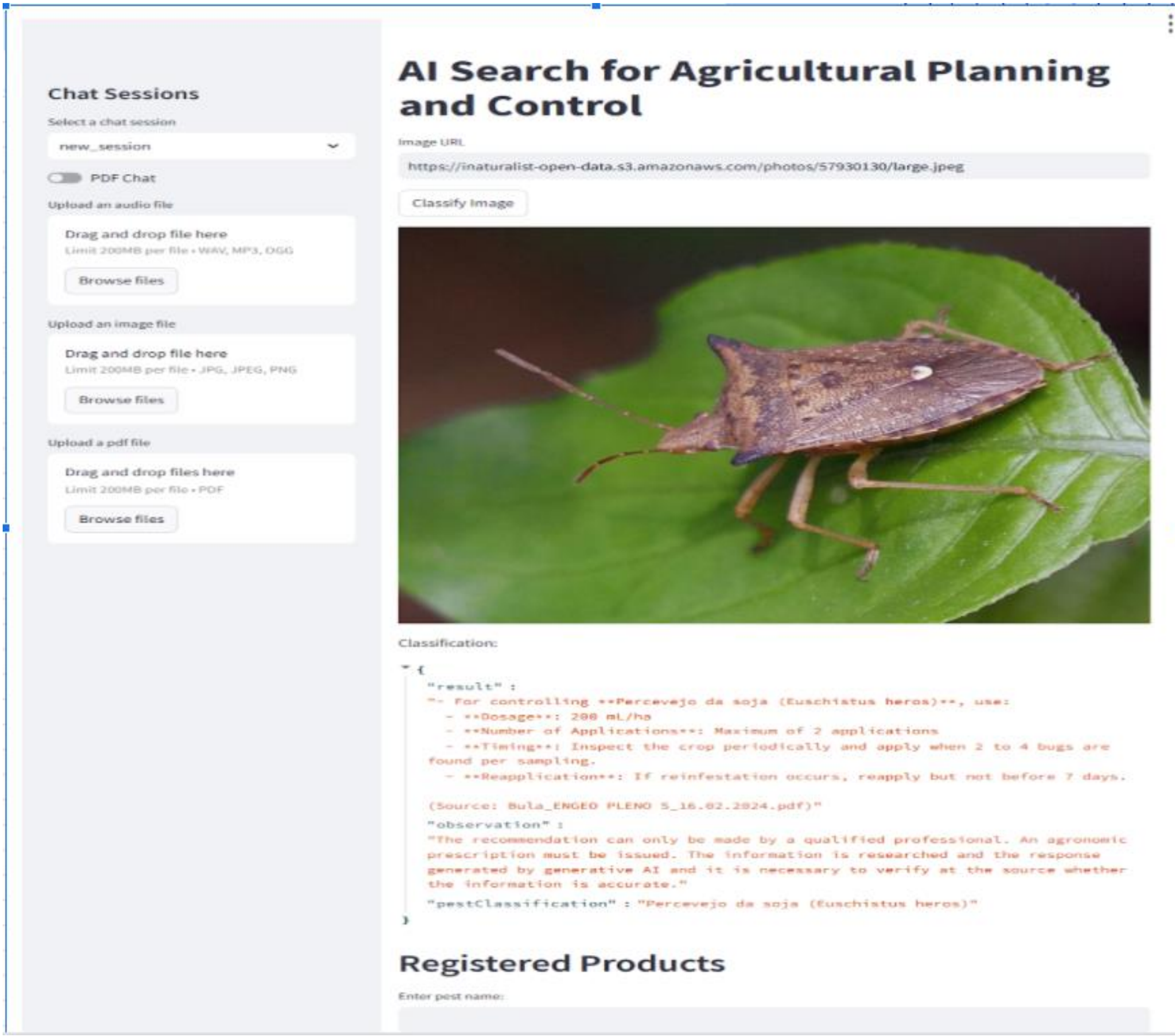
Intensive pesticide use in certain regions has led to genetic selection issues, resulting in pests developing resistance. In this scenario, only the most resilient pests survive, exacerbating the problem.

Tratamento	Dias Após a Aplicação				
	14	21	28	35	42
Testemunha	0,0 C	0,0 C	0,0 D	0,0 D	0,0 E
Boral 500 SC® (500)	89,2 A	77,6 B	51,2 C	31,6 C	20,0 D
Classic® (80)	81,4 B	76,0 B	60,0 C	39,0 C	22,4 D
Classic® (100)	94,6 A	77,6 B	71,0 B	60,0 B	41,6 C
Classic® (120)	97,0 A	79,0 B	73,0 B	60,0 B	40,0 C
Flumyzin 500® (200)	98,0 A	78,0 B	72,0 B	54,0 B	36,6 C
Pivot® (1000)	93,6 A	82,6 B	55,4 C	41,0 C	32,0 C
Imazaquim Nortox® (1000)	93,0 A	91,0 A	83,0 A	72,0 A	63,0 B
Sencor 480® (1000)	96,0 A	92,0 A	84,6 A	68,0 B	68,4 B
Sencor 480® (2000)	94,0 A	92,0 A	86,2 A	82,0 A	78,0 A
Spider® (30)	92,6 A	88,0 A	86,2 A	77,0 A	72,0 A
Spider® (42)	95,4 A	95,0 A	92,0 A	85,0 A	74,0 A
Teste F	214,14**	81,95**	96,28**	30,77**	59,54**
CV (%)	4,87	8,08	8,42	17,75	15,97

The effectiveness of the same herbicide can vary significantly across different weed species. Large Language Models (LLMs) can assist agronomists in quickly identifying the most effective application combinations for optimal weed control.–

Tabela 7.1. Eficiência de alguns herbicidas* aplicados em PPI, pré e pós-emergência, para o controle de plantas daninhas da cultura da soja em solos de Cerrado. Dados compilados da tabela da Comissão de Plantas Daninhas da Região Central do Brasil, 2003.

Nome comum / Nome científico	Bentazon	Bentazon+Acifluorfen	Chlorimuron-ethyl	Clethodim	Cloransulam-methyl	Diclosulam	Dimethenamide	Fenoxaprop-p-ethyl+Clethodim	Fluazifop-p-butyl	Flumetsulan	Flumioxazin PRE	Fomesafen	Fomesafen+Fluazifop[4]	Haloxyfop-R, éster metílico	Imazaquim[5]	Imazethpyr	Lactofen	Metolachlor[6]	Metribuzin	Oxasulfuron	Propaquizafop	Sethoxydim	Tepraloxydin	Trifluralin
Apaga-fogo *Alternathera tenela*	T	-	S	T	-	-	S	-	T	S	S	S	-	-	S	S	S	-	S	-	-	T	-	-
Caruru *Amaranthus viridis*	T	S	S	T	-	S	S	-	T	S	-	S	-	-	S	-	S	-	S	S	-	T	-	S
Carrapicho-rasteiro *Acanthospermum australe*	M	M	S	T	S	S	M	-	T	S	-	M	M	-	S	S	M	T	M	-	-	T	-	T
Mentrasto *Ageratum conyzoides*	S	S	S	T	-	S	S	-	T	S	S	S	S	-	S	M	S	-	S	S	-	T	-	T
Picão-preto *Bidens pilosa*	S	S	S	T	S	S	M	-	T	S	-	S	S	-	S	S	S	S	S	S	-	T	-	T
Falsa-serralha *Emilia sonchifolia*	M	-	S	T	-	-	-	-	-	-	-	S	-	-	M	M	S	-	M	-	-	T	-	T
Erva-de-touro *Tridax procumbens*	S	S	S	T	S	S	-	-	T	S	S	S	S	-	M	-	S	-	-	-	-	T	-	-
Trapoeraba *Commelina benghalensis*	S	S	S	T	M	M	S	-	-	-	-	M	M	-	M	S	S	S	M	-	-	T	-	T
Cordas-de-viola *Ipomoea grandifolia*	S	S	S	T	S	S	T	-	T	M	-	S	S	-	M	S	M	-	M	-	-	T	-	T
Erva-de-santa-luzia *Chamaecyse hirta*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Leiteiro *Euphorbia heterophylla*	T	S	-	T	M	M	T	-	M	M	-	S	S	-	S[1]	S	-	-	T	-	-	T	-	T
Desmodium *Desmodim tortuosum*	T	-	S	T	M	M	T	-	-	-	S	T	-	-	T	T	T	S	S	S	-	T	-	T
Fedegoso *Senna obtusifolia*	T	-	S[2]	T	T	T	M	-	T	S	-	T	-	-	-	T	-	-	-	-	-	T	-	T
Cheirosa *Hyptis suaveolens*	T	S	S[2,3]	T	-	-	S	-	T	S	S	S	S	-	M	S	S	-	M	S	-	T	-	T
Guanxuma *Sida rhombifolia*	S	S	-	T	S	S	T	-	T	S	-	-	-	-	S	S	-	S	S	-	-	T	-	T
Beldroega *Portulaca oleracea*	S	-	S	T	-	-	S	-	T	-	-	S	-	-	S	S	S	-	S	-	-	T	-	M
Poaia-branca *Richardia brasilienses*	-	-	M	T	-	-	-	-	T	S	-	M	S	-	S	M	-	-	S	-	-	T	-	-
Erva-quente *Spermacoce latifólia*	-	-	-	-	-	-	-	-	-	-	S	-	-	-	-	-	S	-	-	-	-	-	-	-
Joá-de-capote *Nicandra physaloides*	S	S	M[2]	T	-	-	S	-	T	T	S	S	S	-	M	S	-	-	S	-	-	T	-	T
Maria pretinha *Solanum americanum*	-	-	T	T	-	-	-	-	T	-	-	-	-	-	S	S	S	S	T	-	-	T	-	T


Chat Sessions
Select a chat session
new_session
PDF Chat
Upload an audio file
Drag and drop file here
Limit 200MB per file • WAV, MP3, OGG
Browse files
Upload an image file
Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG
Browse files
Upload a pdf file
Drag and drop files here
Limit 200MB per file • PDF
Browse files
AI Search for Agricultural Planning and Control
Image URL
https://inaturalist-open-data.s3.amazonaws.com/photos/57930130/large.jpeg
Classify Image
Classification:
{
"result" :
"- For controlling **Percevejo da soja (Euschistus heros)**, use:
- **Dosage**: 200 mL/ha
- **Number of Applications**: Maximum of 2 applications
- **Timing**: Inspect the crop periodically and apply when 2 to 4 bugs are found per sampling.
- **Reapplication**: If reinfestation occurs, reapply but not before 7 days.
(Source: Bula_ENGEO PLENO 5_16.02.2024.pdf)"
"observation" :
"The recommendation can only be made by a qualified professional. An agronomic prescription must be issued. The information is researched and the response generated by generative AI and it is necessary to verify at the source whether the information is accurate."
"pestClassification" : "Percevejo da soja (Euschistus heros)"
}
Registered Products
Enter pest name:

Management and Soil Fertility for the Crop

Military

Douglas de Castilho Gitti
2 Renato Roscoe
3 Lucas de Almeida Rised

Introduction

Proper soil fertility management is an essential condition for the success of soybean crops. Diagnosing the soil's ability to supply nutrients to plants and planning appropriate fertilizer management strategies are important challenges for achieving high productivity and profitability.

Soybean is the most important crop in terms of planted area and profitability for rural producers in Mato Grosso do Sul. Planted on around two million hectares in the state, it is the main crop for most production systems, including integrated ones.

Fertilizer costs make up the largest portion of soybean production costs, representing 25% of the total cost or 36% of the crop cost (Richetti, 2012). In addition to their high cost share, fertilizers are production factors with a strong influence on productivity. Therefore, their correct use can mean the difference between profit or loss.

The proper use of fertilizers involves what has been called "4C Management", which seeks to define the right source, in the right dose, applied at the right time and in the right location (Casarin & Stipp, 2009). These are the foundations of Good Practices for the Efficient Use of Fertilizers (BPUFs), which aim to provide adequate conditions for the balanced supply of nutrients to crops, while minimizing losses (Casarin & Stipp, 2013).

The first challenge is to properly diagnose nutrient deficiencies or excesses, which is done through soil and leaf analysis, assessments of nutrient balances (inputs and outputs), sensors to assess leaf color or chlorophyll levels, or even visual diagnosis of deficiency symptoms. Once the diagnosis is made,

The application encompasses various tasks related to agricultural planning in compliance with Brazilian regulations. Currently, it is developed using Python and C#, integrating an AI language model to assist farmers in creating effective agricultural plans and managing their crops.

A comprehensive database will be established to store Brazilian regulations, dosage limits, and restrictions on pesticide use, which will serve as a resource for the AI language model. This model will be enriched with data tables containing information on seeds, fungicides, herbicides, pesticide efficacy, and soil fertility recommendations derived from both historical documentation and recent reports across different regions.

Additionally, the application features a computer vision model capable of identifying insect infestations and diseases through image analysis. This information is utilized by the language model to recommend effective pesticides that comply with Brazilian standards for pest control.

Furthermore, all searches will provide sources for verification and include warnings regarding the limitations of AI, ensuring users are well-informed.

In 2024, Brazil's agribusiness GDP is projected to reach R$2.50 trillion (approximately $415 billion), with R$1.74 trillion (around $290 billion) attributed to the agricultural sector and R$759.82 billion from the livestock sector. According to the Brazilian Confederation of Agriculture and Livestock (CNA), agribusiness is expected to account for 21.8% of Brazil's GDP in 2024. This significant contribution underscores the vital role of agribusiness in the Brazilian economy, presenting a valuable opportunity for research and the development of applications aimed at enhancing productivity

Assignments

Feature	When users	It should	So that
AI language chat to ask for information on agricultural planning.	Ask a question, the application may interpret this information to identify what the crop is, what kind of infestation there is in this crop and what is the planted area.	The application must return complete information for crop planning. Estimates of fertilizer consumption, agricultural pesticides, crop rotation, seeds appropriate for the region and propose the use of more efficient production technologies.	The application searches a SQL database and a vector database for information and restrictions of local legislation for suggestions. A report is presented in a front-end interface with all suggestions, restrictions and the sources consulted.
Identification of what kind of infestation the crop is suffering based on images of the plants leaves or of the insects that attacked the crop.	Uploads an image of a plant or of an insect.	Identify what type of disease or insect in the image can cause damage to the plant. Quantify through object detection the number of affected plants or number of pests per area to facilitate agronomic recommendations.	The application will identify what infestation the crop is suffering and limit the scope for the search of information. This operation is done by prompt engineering techniques.
Identification of pesticide efficiency based on research report.	Upload and PDF file with the research report.	Identify tables with efficiency data.	The application uses RAG (Retrieved-Augmented Generation) techniques to feed the language model with information about pesticide efficiency.
Identification of more productive seeds according to geographic location and publication of trial results for new cultivars.	Upload a PDF file with the historical documents about pesticides.	Identify productivity, disease resistance, best planting calendar, etc.	The application uses RAG (*Retrieved-Augmented Generation*) techniques to feed the language model with information about the cultivars.

5 Conclusion

This project represents a foundational step towards creating a powerful search engine that leverages the capabilities of LLMs and Azure services. By focusing on responsible AI practices and exploring innovative features, we aim to provide users with valuable insights and enhance their search experience.

Additionally, the extensibility of the project ensures that it can adapt to different languages and regions, broadening its impact and usability.

Goals

High priority	Medium priority	Low priority
Train an Azure AI Vision model to identify the kinds of insects that may infest the crops.	Train an Azure AI Vision model to identify healthy and ill species in the crop.	Insert more images in the Azure Vision training dataset on kinds of insects.
Implement an Azure Document Intelligence model to collect data on research reports.	Implement an Azure Document Intelligence model to collect data on historical documents.	Insert more images in the Azure Vision training dataset on healthy and ill species.
Implement a language model to answer questions about agricultural planning based on information of crops, infestations, etc. given by the user.	Implement prompt engineering to limit the scope of the model based on images loaded and classified by the application.	
Include a RAG to the language model to incorporate information from recent research on pesticide efficiency.	Include a RAG to the language model to incorporate information from historical documents on pesticide efficiency.	
Implement a front-end interface for the application and improve the backend.		

Research Overview

1. The theme of this literature review focuses on soybean cultivation, particularly in the context of agricultural practices, pest management, and the impact of environmental factors. The purpose is to synthesize existing research to understand the challenges and advancements in soybean production. This is significant as soybean is a major crop globally, contributing to food security and economic stability.

Literature Analysis

- **Theoretical Framework**: This research is grounded in agroecology and integrated pest management. These frameworks emphasize sustainable agricultural practices that balance productivity with environmental health.

Key Findings:

- **Pest Management**: The literature highlights the increasing complexity of pest management in soybean crops due to the emergence of herbicide-resistant weeds. Studies indicate that effective management requires a combination of cultural practices and herbicide applications.
- **Environmental Impact**: Research shows that climatic conditions significantly affect soybean growth phases. Optimal sowing times are crucial for maximizing yield and minimizing pest outbreaks.
- **Soil Health**: The importance of soil management practices, including the use of cover crops and crop rotation, is emphasized to enhance soil fertility and reduce pest pressures.
- **Methodological Evaluation**: Most studies utilize experimental designs with controlled variables to assess the effectiveness of various treatments. While these methods provide robust data, they often lack long-term ecological assessments, which are crucial for understanding sustainable practices.

Research Gaps and Future Directions

- **Research Gaps**: There is a notable gap in longitudinal studies that examine the long-term effects of integrated pest management strategies on soybean yield and soil health. Additionally, the impact of climate change on pest dynamics in soybean production remains underexplored.
- **Future Research Recommendations**: Future research should focus on developing adaptive management strategies that consider climate variability. Incorporating farmer participatory approaches could enhance the relevance and applicability of findings.

Discussion

- **Research Contributions**: This literature review contributes to the field by consolidating knowledge on soybean cultivation practices and highlighting the interplay between pest management and environmental factors.
- **Theoretical and Practical Implications**: The findings have significant implications for both theory and practice, suggesting that a holistic approach to soybean cultivation can lead to improved sustainability and productivity.

Conclusion

- **Summary**: This review underscores the critical role of integrated practices in soybean cultivation, revealing the complexities of pest management and the necessity for adaptive strategies.
- **Limitations**: The limitations of this literature review include a potential bias towards recent studies and a lack of comprehensive data on global soybean production practices.

References

- Embrapa. (2013). *Soil Management and Fertility for Soybean Cultivation*.
- Broch, D.L., & Ranno, S.K. (2012). *Soybean Production Technologies*.