UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DE COMPUTAÇÃO

LADSON GOMES SILVA

Applying Generative AI to Plant Disease Diagnosis: A Multimodal Agent for Supporting Smallholder Farmers

Recife

2025

LADSON GOMES SILVA

Applying Generative AI to Plant Disease Diagnosis: A Multimodal Agent for Supporting Smallholder Farmers

The dissertation was submitted to the Graduate Program in Computer Science at the Informatics Center of the Federal University of Pernambuco as a partial requirement for obtaining a Master's degree in Computer Science.

Area of Concentration: Computational Intelligence

Advisor: Prof. Dr. Stefan Michael Blawid

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Silva, Ladson Gomes.
Applying generative ai to plant disease diagnosis: a multimodal agent for supporting smallholder farmers / Ladson Gomes Silva. - Recife, 2025.
130 f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Informática, Programa de Pós-Graduação em Ciência da Computação, 2025.
Orientação: Stefan Michael Blawid.
Inclui referências e apêndices.

1. Generative artificial intelligence; 2. Retrieval-augmented generation; 3. Plant disease diagnosis; 4. Conversational systems; 5. Knowledge retrieval; 6. Agricultural technology. I. Blawid, Stefan Michael. II. Título.

UFPE-Biblioteca Central

Ladson Gomes Silva

“Applying Generative AI to Plant Disease Diagnosis: A Multimodal Agent for Supporting Smallholder Farmers”

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Inteligência Computacional.

Aprovado em: 28/07/2025.

BANCA EXAMINADORA

_____

Prof. Dr. Filipe Carlos de Albuquerque Calegario
Centro de Informática / UFPE

_____

Prof. Dr. , André Angelo Medeiros Gomes
Departamento de Agronomia / UFRPE

_____

Prof. Dr. Stefan Michael Blawid
Centro de Informática / UFPE
(**orientador**)

AGRADECIMENTOS

Ao meus pais, Cleonice Gomes e Luiz Gonzaga Gomes, por ter me proporcionado educação e incentivo constante para investir na minha formação acadêmica. Eles foram os primeiros da família a conhecer o poder transformador social da educação, e sempre nos incentivaram a isso.

Ao meu irmão e minha cunhada, Layrton e Marcia, por me apoiarem e incentivarem a continuar sempre me esforçando para alcançar meus objetivos e ir além.

Aos meus amigos, Ilária Martina e Ruam Pastor, que dividiram apartamento comigo por grande parte dessa jornada, e foram minha inspiração mais próxima no dia a dia, de seguir com o mestrado.

Aos meus amigos que são muitos para citar, mais em especial, Igor e Laís, que em vários momentos me arrastaram para conseguir seguir com o mestrado, nos apoiando mutuamente e nos ajudando a superar os desafios.

Ao meu professor, Stefan Blawid, por me guiar e me apoiar em todas as etapas do mestrado. Seu apoio nas minhas ideias, por vezes distantes da sua zona de conforto, foi o que permitiu desenvolver esse trabalho. Obrigado por me dar força e incentivo para continuar, mesmo quando a situação parecia desesperada.

Aos desenvolvedores da OpenAI e Anthropic, por terem desenvolvido os modelos que foram tanto o objeto de estudo do meu trabalho, quanto minhas ferramentas para auxiliar na escrita e revisão do texto desse trabalho quanto a exaustão da jornada dupla de trabalho não me permitia avançar em tempo hábil.

Por fim, mas não menos importante, um agradecimento a UFPE, onde estive por um terço da minha vida, e onde aprendi lições valiosas que contribuíram no desenvolvimento e formação de minha vida pessoal, acadêmica e profissional.

”The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say”
- J.R.R. Tolkien, The Fellowship of the Ring

RESUMO

O presente trabalho busca entender como técnicas mais recentes de inteligência artificial generativa—como os Modelos de Linguagem de Grande Escala (LLMs) e a Geração Aumentada por Recuperação (RAG) podem ser aplicadas no diagnóstico de doenças em plantas. O estudo envolve a criação do LIMMO, um agente conversacional multimodal pensado para ajudar agricultores familiares por meio de conversas em linguagem natural e também pelo envio de imagens. Combinando modelos avançados de linguagem, análise de imagem com redes neurais e recuperação de informações especializadas, o sistema tenta lidar com alguns dos principais desafios do campo, como a falta de acesso a dados confiáveis e a dificuldade de conseguir apoio técnico em tempo real. A arquitetura final do sistema integra múltiplas fontes de conhecimento, incluindo uma base de dados vetoriais local, APIs de busca na web (como a Tavily) e a API da Embrapa para acesso a dados agrícolas especializados. Para análise de imagens, o sistema utiliza uma abordagem redundante, com a API CultivAI como método primário e o processamento baseado em GPT como sistema de backup quando a extração principal falha ou é questionada pelo usuário. Esta arquitetura modular com Protocolos de Contexto de Modelo (MCPs) demonstrou ser mais eficiente que as implementações anteriores baseadas em sistemas multi-agentes. A metodologia de avaliação utilizou 100 perguntas sintéticas, analisando precisão, consistência factual, qualidade da recuperação e utilidade das respostas. Os resultados mostram que o sistema RAG com acesso a fontes externas supera significativamente as abordagens que dependem apenas de conhecimento local, particularmente em consultas complexas ou fora do escopo imediato da base de conhecimento. Para o futuro, o trabalho aponta caminhos que incluem expansão das fontes de dados, testes massivos em condições reais, desenvolvimento de um sistema de rastreamento de doenças, e mecanismos para diferenciação de ferramentas generalistas como ChatGPT, Gemini ou Perplexity AI em consultas fora do escopo especializado.

Palavras-chaves: Inteligência Artificial Generativa; Modelos de Linguagem de Grande Escala; Geração Aumentada por Recuperação; Diagnóstico de Doenças em Plantas; Agente de IA Multimodal; Agricultura Familiar; Tecnologia Agrícola; Análise de Imagens; Recuperação de Conhecimento; Sistemas Conversacionais

ABSTRACT

This research examines how recent advances in generative artificial intelligence, particularly Large Language Models and Retrieval-Augmented Generation (RAG), can be applied to plant disease diagnosis. It introduces LIMMO, a multimodal conversational agent designed to assist smallholder farmers through natural language conversations and image-based interactions. By combining modern language models, image analysis using deep learning, and smart information retrieval from specialized sources, the system addresses key challenges in agricultural environments, such as limited access to technical support and reliable data.

The final system architecture integrates multiple knowledge sources, including a local vector database, web search capabilities through the Tavily API, and specialized agricultural data from the Embrapa API. For image analysis, the system employs a redundant approach, using the CultivAI API as the primary method while seamlessly falling back to GPT-based processing when primary extraction fails or is questioned by the user. This modular architecture with specialized Model Context Protocols (MCPs) proved more efficient than earlier implementations based on multi-agent systems.

The evaluation methodology utilized 100 synthetic questions, analyzing accuracy, factual consistency, retrieval quality, and response utility. Results demonstrate that the RAG system with access to external sources significantly outperforms approaches relying solely on local knowledge, particularly for complex queries or those outside the immediate scope of the knowledge base. The dissertation concludes by outlining future directions, including expanding data sources, conducting large-scale real-world testing, developing a disease tracking system, and creating mechanisms to differentiate from generalist tools like ChatGPT, Gemini, or Perplexity AI when handling queries outside the specialized scope.

Keywords: Generative Artificial Intelligence; Large Language Models; Retrieval-Augmented Generation; Plant Disease Diagnosis; Multimodal AI Agent; Smallholder Farmers; Agricultural Technology; Image Analysis; Knowledge Retrieval; Conversational Systems

LIST OF FIGURES

LIST OF FRAMES

LIST OF TABLES

LISTA DE ABREVIATURAS E SIGLAS

AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
BERT	Bidirectional Encoder Representations from Transformers
CNN	Convolutional Neural Network
CoT	Chain-of-Thoughts
ELISA	Enzyme-Linked Immunosorbent Assay
FAISS	Facebook AI Similarity Search
GNN	Graph Neural Networks
GPT	Generative Pre-trained Transformer
LAMP	Loop-Mediated Isothermal Amplification
LLM	Large Language Model
LRM	Large Reasoning Model
LSTM	Long Short-Term Memory
MCP	Model Context Protocol
MLP	Multi-Layer Perceptrons
NLP	Natural Language Processing
OCR	Optical Character Recognition
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
RAG	Retrieval-Augmented Generation
RAGAS	Retrieval-Augmented Generation Assessment Suite
RNN	Recurrent Neural Network

CONTENTS

1 INTRODUCTION

1.1 MOTIVATION

Agriculture remains the foundation of global food security, yet it is constantly threatened by plant diseases and pests. In Brazil, one of the world's agricultural powerhouses, the impact of these challenges is particularly severe, affecting smallholder farmers who produce over 70% of the nation's food supply (IBGE, 2017). Despite their vital role, these farmers often lack timely access to reliable plant health diagnostics, leading to delayed interventions, increased use of pesticides, and substantial economic losses.

At the same time, the rapid advances in AI, particularly in Generative AI and Large Language Models (LLMs), have opened unprecedented opportunities to bridge this knowledge gap. The emergence of Retrieval-Augmented Generation (RAG) frameworks, combined with deep learning models for image analysis, enables the development of systems capable of understanding natural language queries, analyzing visual symptoms, and providing expert-level guidance almost instantly. These capabilities can empower farmers and agronomists to make faster, more informed decisions, potentially transforming agricultural diagnostics from a reactive to a proactive process.

While traditional AI approaches in agriculture have focused on either image recognition or text-based interactions separately, the integration of multimodal capabilities through Model Context Protocols (MCPs) offers a promising path forward. By combining the visual analysis strengths of Convolutional Neural Networks (CNNs) with the contextual understanding of LLMs, modern systems can now mimic the holistic diagnostic approach of human experts, considering multiple sources of evidence before reaching conclusions.

However, the adoption of AI in agriculture faces its own set of challenges. Diagnostic tools must handle diverse field conditions, interpret multimodal data accurately, and deliver interpretable, actionable insights that earn user trust. This dissertation is driven by the vision of creating an accessible, robust, and intelligent assistant capable of addressing these challenges and extending expert support to those who need it most.

1.2 PROBLEM STATEMENT

While image-based plant disease recognition and AI-driven advisory tools have shown promising results in controlled environments, several critical gaps remain before they can be deployed effectively in real-world agricultural contexts:

1. **Narrow Approaches to Diagnosis:** Current systems often focus on narrow tasks, such as single-disease classification or symptom detection, neglecting the complex reasoning and context integration required for accurate diagnosis and treatment recommendations. These narrow approaches miss opportunities for cross-modal validation and complementary analysis.

2. **Limited Adaptability to Field Conditions:** Many models are trained on laboratory-grade images or curated datasets that do not reflect the variability in lighting, backgrounds, and symptom manifestations encountered in actual field scenarios. The resulting performance gap undermines their practical utility precisely where they are needed most.

3. **Lack of Integration Between Modalities:** Few solutions seamlessly combine text-based symptom descriptions with image analysis, limiting their ability to reflect the real communication flow between farmers and experts. This modality gap represents a missed opportunity to leverage complementary information sources for more robust diagnostics.

4. **Hallucinations and Factual Inconsistency:** Generative AI systems, while powerful, are prone to producing plausible-sounding but factually incorrect information, a critical risk in agricultural contexts where incorrect advice can lead to significant economic and environmental consequences.

5. **User Experience:** Existing tools are often designed for researchers or technical users, failing to consider the usability requirements of farmers working in diverse and resource-limited environments. The resulting adoption barriers limit the real-world impact of potentially valuable technologies.

By addressing these challenges, this research seeks to design, implement, and evaluate a conversational AI system that integrates RAG, multimodal inputs, and a

modular architecture to provide reliable, explainable, and user-friendly plant disease diagnostics.

1.3 OBJECTIVES

The main objective of this dissertation is to design, implement, and evaluate a prototype of an intelligent conversational agent capable of assisting in the diagnosis of plant diseases through multimodal interactions. To achieve this overarching goal, the following specific objectives are pursued:

1. **Design and Implement a Hybrid RAG-Enhanced Architecture:** Develop an iterative prototyping methodology to explore different architectural approaches for plant disease diagnostics, culminating in a streamlined single-agent design enhanced with retrieval-augmented generation.
2. **Enable Multimodal Input Processing and Analysis:** Create a system capable of accepting and analyzing text, image, and audio inputs, with specialized modular components (MCPs) handling each modality while maintaining a coherent user experience.
3. **Combine Convolutional Neural Networks with RAG:** Integrate deep learning-based image analysis with retrieval-grounded language generation, by extracting features from images and using them as input to the RAG, creating a diagnostic process that leverages both visual pattern recognition and domain-specific knowledge.
4. **Optimize for Practical Performance Metrics:** Balance technical sophistication with practical considerations like response time, maintenance complexity, and diagnostic accuracy to ensure the system meets real-world agricultural needs.
5. **Evaluate System Effectiveness and User Experience:** Assess the system through both quantitative benchmarks (accuracy, response time) and qualitative feedback from domain experts, focusing on diagnostic reliability and interface usability.

By addressing these objectives, this dissertation aims to contribute a robust, user-centered solution that brings advanced AI capabilities to the field, supporting farmers and agricultural professionals in making more timely and informed decisions.

1.4 CONTRIBUTIONS

This dissertation offers several key contributions to the fields of agricultural diagnostics, computer vision, and conversational AI:

1. **A Novel Modular Architecture for Plant Disease Diagnostics:** The development of a streamlined, single-agent architecture augmented with RAG and modular components (MCPs) that balances technical sophistication with operational maintainability, representing a significant advancement over both simple RAG workflows and complex multi-agent systems.
2. **Systematic Comparison of Architectural Approaches:** An empirical evaluation of different system designs, from basic RAG pipelines to multi-agent orchestration and finally to the optimized hybrid approach, providing valuable insights on the trade-offs between complexity, performance, and maintainability in agricultural AI systems.
3. **Integration of CNN-Based Visual Analysis with RAG:** A practical implementation demonstrating how CNNs for image analysis can be effectively combined with retrieval-augmented generation, reducing hallucinations while maintaining response quality and diagnostic accuracy.
4. **A Multimodal Agricultural Diagnostic System:** The creation of a functional prototype that processes text, image, and audio inputs through a unified framework, closely mimicking the natural diagnostic workflow of human agronomists while maintaining system coherence.
5. **Empirical Insights on Performance and Usability:** Quantitative and qualitative evaluations demonstrating the progressive improvements in diagnostic accuracy (from 60% to over 90%), response times, and user satisfaction across system iterations.

Together, these contributions advance the state of the art in applying AI for plant health monitoring, bridging technical innovations and practical needs to empower farmers and agricultural professionals.

1.5 STRUCTURE OF THE DISSERTATION

This dissertation is structured as follows:

- **Chapter 1** introduces the context, motivations, and objectives of this work.
- **Chapter 2** presents a historical overview of the evolution of Generative AI tools and their impact on research and prototyping.
- **Chapter 3** reviews related work in LLM applications for agriculture, plant disease diagnosis using deep learning, and AI-powered diagnostic systems.
- **Chapter 4** describes the proposed methodology, including the system design, model training, and technologies used.
- **Chapter 5** presents and discusses the results of the developed prototype.
- **Chapter 6** concludes the dissertation, summarizing key findings and proposing directions for future research.

2 THE EVOLUTION OF GENERATIVE AI AND NATURAL LANGUAGE INTERFACES

2.1 INTRODUCTION

AI has rapidly evolved over the past decades, fundamentally transforming the way we process and interact with information. From early rule-based systems to modern deep learning techniques, AI has expanded its reach across numerous domains, including healthcare, finance, and agriculture. One of the most impactful advancements has been the emergence of models capable of understanding and generating natural language, enabling machines to communicate with humans in increasingly intuitive ways.

For the agricultural sector, these advancements offer unique opportunities. Farmers and agronomists can now access intelligent systems that help identify plant diseases, suggest treatments, and provide real-time decision support, all through simple conversations in natural language or by sending images. However, to appreciate the capabilities and limitations of these modern systems, it is essential to understand the technological foundations that made them possible.

This chapter introduces the core concepts of AI and machine learning, briefly discusses the evolution of neural networks, and explains how the introduction of transformers and LLMs revolutionized natural language understanding and generation. Finally, it highlights how these innovations set the stage for the development of conversational diagnostic systems in agriculture.

2.2 NEURAL NETWORKS: A BRIEF FOUNDATION

Before discussing transformers, it is important to review the neural network architectures that paved the way for them. Artificial Neural Networks (ANNs), such as Multi-Layer Perceptrons (MLPs), consist of interconnected layers of neurons trained using gradient descent and backpropagation. While MLPs form the conceptual basis for all subsequent deep learning models, they lack built-in mechanisms to handle structured data like sequences or images effectively.

Convolutional Neural Networks introduced convolutional filters to capture local and

hierarchical patterns in spatial data, revolutionizing computer vision tasks. For sequential data, Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks, were developed to model temporal dependencies by incorporating feedback loops. However, RNNs face major challenges, including vanishing and exploding gradients, which make learning long-range dependencies difficult (PASCANU; MIKOLOV; BENGIO, 2013). Although LSTMs mitigate some of these issues through gating mechanisms (HOCHREITER; SCHMIDHUBER, 1997), their sequential nature inherently limits parallelization and efficiency.

More recently, architectures such as Vision Transformers (ViTs) have demonstrated the versatility of transformer-based models, extending their use beyond text to computer vision applications (DOSOVITSKIY et al., 2021).

2.3 THE TRANSFORMER REVOLUTION

Transformers, introduced by Vaswani et al. (2017), addressed the limitations of RNNs by eliminating recurrence and relying entirely on a self-attention mechanism. At the heart of the transformer lies the **Scaled Dot-Product Attention** mechanism. Given a set of queries Q, keys K, and values V, the attention function computes an output as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V$$

Here, d_k is the dimensionality of the key vectors. The computation proceeds in three main steps:

1. **Similarity Calculation:** The dot product between each query and all keys ($QK^\top$) measures the similarity or compatibility between the query and each key.

2. **Scaling:** The result is divided by $\sqrt{d_k}$ to prevent the dot products from growing too large in magnitude. Without this scaling—especially when d_k is large—the softmax function can push the outputs into regions with extremely small gradients, making training unstable or slow.

3. **Weighting Values:** A softmax is applied to obtain a probability distribution over the keys, and the output is computed as a weighted sum of the values V, with the

weights derived from this distribution.

This mechanism enables the model to capture long-range dependencies while supporting full parallelization during training, unlike RNNs which process sequences step by step.

Moreover, multi-head attention allows the model to attend to different representation subspaces simultaneously, enhancing its expressiveness (VASWANI et al., 2017). Empirical results showed that transformers significantly outperformed RNNs in both accuracy and training efficiency on various tasks, including machine translation, marking a pivotal moment in deep learning research.

2.4 IMPLICATIONS FOR AGRICULTURAL DIAGNOSTIC SYSTEMS

These architectural advantages directly support the development of agricultural diagnostic systems:

- **Multimodal integration**: Transformers can jointly process textual descriptions and visual data, ideal for conversational plant health diagnostics.
- **Transfer learning**: Models pretrained on massive datasets can be fine-tuned for domain-specific tasks such as plant disease recognition or agronomic Q&A.
- **Explainability**: Attention outputs can be visualized, offering transparency into the model's reasoning—crucial for user trust among agronomists.
- **Inference efficiency**: Due to parallelizable architecture, transformers enable real-time interaction even in resource-limited field environments.

The system developed in this dissertation builds on these advantages, integrating multimodal capabilities and explainability to empower field diagnostics and decision-making in real-world agricultural contexts.

2.5 FROM GPT-2 TO GEMINI: A TIMELINE OF MODERN LLMS

Recent years have witnessed a rapid and transformative evolution in large language models (LLMs). Beginning with GPT-2 in 2019, which impressed the research commu-

nity with its fluent text generation, the development of LLMs has advanced toward increasingly capable and versatile multimodal models such as GPT-4 and Gemini. This progress has involved not only scaling models to unprecedented sizes but also optimizing smaller, more efficient models.

Transformer-based models like BERT (2018) introduced bidirectional language understanding, achieving state-of-the-art results on numerous Natural Language Processing (NLP) tasks through pre-training on large text corpora and fine-tuning for specific applications (DEVLIN et al., 2019). Around the same time, OpenAI's GPT series demonstrated the power of unidirectional generative pre-training. GPT-2 (2019), with 1.5 billion parameters, set new standards for fluent text generation, while GPT-3 (2020), with 175 billion parameters, showcased impressive few-shot learning capabilities, enabling tasks such as translation, question answering, and arithmetic from minimal examples. These milestones highlighted that scaling data and parameters could produce emergent capabilities.

Transformer-based models diverged early in how they learn from text. BERT introduced deep bidirectional contextual representations: during pre-training, BERT jointly conditions on both left and right context in all layers, enabling the model to encode complete sentence semantics (DEVLIN et al., 2018). Such bidirectional encoding makes BERT and similar models well suited for language understanding tasks such as sentiment analysis, question answering, and named-entity recognition, where full-sentence comprehension is essential. In contrast, GPT models adopt a unidirectional (autoregressive) training objective, factorizing the likelihood of a sequence into a product of conditional probabilities and predicting each token based only on preceding tokens (DEVLIN et al., 2018). This causal modeling aligns naturally with text generation, allowing GPT models to excel at composing coherent stories, dialogue, or code. Despite the success of bidirectional models like BERT, later research notes that GPT-style unidirectional models still achieve state-of-the-art performance by scaling model size and training data (DEVLIN et al., 2018). Comparative studies further observe that fully unidirectional models (GPT) and fully bidirectional models (BERT) sit at opposite ends of a spectrum, and hybrid approaches attempt to combine bidirectional context and attention to balance generation and understanding (ARTETXE et al., 2022). Understanding these foundational differences clarifies why GPT-type models dominate generative applications while BERT-type models remain the backbone of modern language com-

prehension systems.

By 2022-2023, the development of LLMs began to diverge into two main trends: the creation of ever-larger models and the design of efficient, specialized smaller models. OpenAI's GPT-4 (2023) introduced multimodality, handling both text and images, and achieved human-level performance on many academic and professional benchmarks. Anthropic's Claude (2023) emphasized helpfulness and safety, using a "Constitutional AI" alignment strategy that relies on guiding principles rather than purely human feedback. Google DeepMind's Gemini, announced in late 2023 as a successor to PaLM 2, represented a family of multimodal models capable of processing text, images, audio, code, and more. Gemini integrated advanced reasoning capabilities and was positioned as a strong competitor to GPT-4.

Meta's LLaMA models, released starting in mid-2023, demonstrated that smaller open-source models (7B-70B parameters) could achieve competitive performance when trained on high-quality data. Their permissive licenses fueled widespread community adaptation and fine-tuning. Notably, smaller models began achieving remarkable results: Microsoft's Phi-1 (1.3B) and Phi-2 (2.7B) models, trained on carefully curated high-quality data, performed comparably or even surpassed much larger models on various benchmarks. Phi-2, for instance, exhibited outstanding reasoning and language understanding, outperforming models with more than 13 billion parameters.

Similarly, Mistral AI's 7B model (2023) outperformed larger models such as LLaMA-2 13B and even 34B on tasks involving reasoning, mathematics, and code, benefiting from architectural optimizations like grouped-query attention. Google's Gemma models, introduced in early 2024, epitomized the trend toward open, efficient LLMs. Gemma encompasses a family of open-access models (ranging from 2B to 7B parameters) incorporating innovations from Gemini, and supports text, code, and vision tasks, enabling competitive performance on modest hardware.

Community-driven initiatives, such as TinyLlama, illustrate the "small but powerful" movement. TinyLlama aims to pre-train a 1.1B-parameter model on an unprecedented 3 trillion tokens to achieve robust performance at a tiny scale.

- **2018:** BERT (DEVLIN et al., 2019) introduces bidirectional transformers for language understanding; GPT-1 (117M) demonstrates generative pre-training.
- **2019:** GPT-2 (1.5B) showcases high-quality text generation; Google introduces

T5 and XLNet with alternative pre-training objectives.

- **2020:** GPT-3 (175B) pioneers few-shot learning at scale.
- **2021:** Research on scaling laws (e.g., Chinchilla by DeepMind) emphasizes the importance of data quality and training efficiency over sheer parameter count.
- **2022:** Google's PaLM (540B) and OpenAI's Codex (a GPT-3 variant for code) expand capabilities; open-source large models like BLOOM and OPT emerge. ChatGPT (GPT-3.5) popularizes conversational AI.
- **2023:** GPT-4 (multimodal) sets new benchmarks. Meta releases LLaMA (7B-65B), democratizing access to advanced LLMs. Anthropic launches Claude v1 and v2. Microsoft's Phi models demonstrate high efficiency at smaller scales. Mistral-7B surpasses larger models in key benchmarks. OpenAI begins deployment of inference-optimized GPT-4 variant internally known as **o1**.
- **2024:** Google DeepMind introduces Gemini (multimodal, multiple scales) and releases Gemma (open 2B/7B models). Anthropic advances Claude Next. Research intensifies on ultra-efficient training (e.g., TinyLlama) and specialized domain-focused LLMs. OpenAI continues iterative updates with improved variants **o2**, enhancing performance and cost efficiency of GPT-4 Turbo deployments.
- **2025:** DeepSeek releases DeepSeek-R1, an open-weight reasoning model (with distilled smaller models) that claims performance comparable to top models like OpenAI's o1, at a much lower cost. Its emergence causes renewed attention to cost-efficiency, reasoning-centric RL methods, and open-weight LLMs. OpenAI launches **GPT-4o (o3)**, a unified multimodal model trained natively on text, vision, and audio, further blurring modality boundaries in LLM interaction. DeepSeek-Prover targets formal theorem proving. Claude 3 and Gemini 2 expand into native multimodality.

This timeline demonstrates how the development of LLMs has advanced rapidly, balancing scale with efficiency and accessibility. The growing focus on multimodal and open-source models reflects a shift toward specialized, hybrid capabilities, paving the way for domain-specific deployments and real-time applications on diverse hardware platforms.

2.6 BENCHMARKING PROGRESS: MMLU, BIG-BENCH, AND BEYOND

The rapid advancement of large language models has driven the development of comprehensive benchmarks to evaluate their breadth and depth of capabilities. These benchmarks serve as critical tools to assess general knowledge, reasoning, dialogue skills, and ethical considerations, providing a standardized basis for comparing models across different tasks and settings.

- **MMLU (Massive Multitask Language Understanding)** (HENDRYCKS et al., 2021): MMLU evaluates accuracy across 57 diverse academic and professional subjects, including history, mathematics, science, and law. Models are tested in zero-shot or few-shot settings using multiple-choice questions, offering a robust measure of the factual and disciplinary knowledge captured during pre-training. While GPT-3 achieved above-chance performance on many subjects, newer models such as GPT-4 and Claude have approached or surpassed human-level accuracy in several areas. MMLU has thus become a de facto standard for assessing general knowledge and reasoning in LLMs, often summarized as a single aggregate score.

- **BIG-Bench (Beyond the Imitation Game Benchmark)** (SRIVASTAVA et al., 2023): BIG-Bench is a collaborative benchmark comprising over 200 diverse tasks contributed by the research community to probe LLMs' generalization and reasoning abilities beyond conventional NLP challenges. It includes traditional NLP tasks as well as creative and novel challenges, such as logic puzzles, common-sense reasoning, code generation, and inventive language use. Performance is analyzed as a function of model scale, revealing that larger models, such as GPT-4, excel across most tasks, while smaller models often struggle with more complex or abstract challenges.

- **MT-Bench (Multi-Turn Benchmark)** (ZHENG et al., 2023): MT-Bench is designed to assess the quality of multi-turn dialogue in conversational agents. Developed by the Vicuna team, it consists of open-ended questions that require models to engage in extended interactions for clarification and elaboration. The benchmark evaluates the ability to follow intricate instructions, maintain context over multiple

exchanges, and generate helpful, accurate responses. To ensure scalability and consistency, MT-Bench employs an LLM-as-a-judge approach, using GPT-4 as the evaluator. This method has shown over 80% agreement with human judgments, enabling efficient and reliable comparison of conversational abilities.

- **HELM (Holistic Evaluation of Language Models)** (LIANG et al., 2022): HELM offers a comprehensive, multidimensional evaluation framework that goes beyond simple accuracy metrics. It assesses models across various tasks—such as summarization, dialogue, and reading comprehension—and evaluates factors like calibration, robustness, fairness, toxicity, bias, and efficiency. By providing a detailed performance profile rather than a single score, HELM encourages a nuanced understanding of model capabilities and limitations. Regular updates to HELM ensure that the community can monitor progress not only in raw performance but also in ethical and practical aspects relevant to real-world deployment.

Together, these benchmarks inform both the development and deployment of LLMs, guiding trade-offs between accuracy, safety, and usability. Their continued evolution plays a role in shaping the responsible and effective integration of language models across diverse domains.

2.7 FROM CHAIN-OF-THOUGHT TO REASONING MODELS: HOW OUR NOTION OF “INTELLIGENCE” IN LLMS KEEPS SHIFTING

2.7.1 Why this matters

The previous sections have outlined how larger models, richer pre-training corpora, and retrieval or tool augmentation have steadily improved benchmark performance (section 2.6). However, in parallel, the *criteria* we use to define an LLM as “intelligent” have evolved just as rapidly. Early gains in multiple-choice benchmarks (e.g., MMLU) once appeared impressive, but more challenging suites such as HLE and BIG-BENCH-EH soon exposed brittle shortcut behaviors (HENDRYCKS et al., 2020; PHAN et al., 2025; KAZEMI et al., 2025). This realization has driven two intertwined shifts:

(i) a methodological shift toward *visible reasoning*, involving chain-of-thought prompts, self-reflection, and explicit tool use;

(ii) a conceptual shift from viewing models as opaque pattern-matchers to considering them as emergent “reasoners.”

2.7.2 The chain-of-thought era

Early experiments revealed that even GPT-3 could solve complex arithmetic problems when encouraged to “think step by step.” This insight was quickly formalized through techniques like chain-of-thought prompting (CoT), REACT, and self-reflection pipelines. These developments gave rise to benchmarks that evaluate a model’s reasoning process rather than just final answers (section 2.6). Consequently, the field’s implicit definition of intelligence broadened: models were now expected not only to provide correct answers but also to *expose* their reasoning and maintain internal consistency.

2.7.3 Large Reasoning Models

Between 2024 and 2025, purpose-built “reasoning” variants emerged, including OpenAI’s o- series, DeepSeek-R1, Gemini–Thinking, and Claude 3.7 Sonnet Thinking. These models introduce the notion of a dedicated “thinking budget” during inference. Apple’s recent *Illusion of Thinking* study (SHOJAEE et al., 2025) provides a systematic examination of these models’ internal processes. In controlled puzzle environments, the study identified three distinct regimes:

1. **Low complexity:** Standard LLMs often outperform Large Reasoning Models (LRMs), using fewer tokens and achieving higher efficiency.
2. **Medium complexity:** LRMs begin to excel as additional reflection offsets planning costs, improving success rates.
3. **High complexity:** Both standard LLMs and LRMs fail, with LRMs intriguingly exhibiting reduced reasoning effort as task difficulty increases, suggesting a potential scaling limit during inference.

These findings align with insights from the agent-centric discussion in section 2.8.4, where new process-supervision datasets such as PROCESSBENCH and BFCL-V2 emerged

to differentiate between “rote” and “reasoned” solutions (ZHENG et al., 2024; YAN et al., 2024). This shift underscores the limitations of traditional accuracy-focused metrics and highlights the need for process-oriented evaluation.

2.7.4 So what counts as intelligence now?

Bringing these strands together reveals an evolving definition of intelligence:

- **External behavior**: Still essential, as benchmark performance remains a core indicator (section 2.6), but no longer sufficient on its own.
- **Transparency of process**: Models must be able to reveal and verify their internal reasoning steps; agentic evaluation frameworks, such as AGENT-AS-A-JUDGE, push in this direction (ZHUGE et al., 2024).
- **Robust scalability**: Truly intelligent systems should degrade *gracefully* as task difficulty increases. Apple’s study (SHOJAEE et al., 2025) demonstrates that current LRMs fail to meet this criterion, indicating an important frontier for future research.

2.8 ECOSYSTEM OF TOOLS AND DEVELOPMENT PARADIGMS

The rise of LLMs has given birth to a rich ecosystem of frameworks and design paradigms that support the development of increasingly sophisticated applications.

2.8.1 LangChain and LangGraph

LangChain (LANGCHAIN, 2023) has emerged as a widely adopted framework that abstracts the complexities inherent in building LLM-driven applications. It offers a modular architecture for chaining prompts, models, and external tools into multi-step pipelines, enabling the creation of advanced systems such as retrieval-augmented generation (RAG), chatbots, and autonomous agents. By providing standardized interfaces to various LLM APIs and data sources, LangChain facilitates seamless integration with different models and vector databases, requiring minimal code modifications. This design paradigm not only accelerates prototyping but also promotes reproducibility and scalability in both research and industry contexts.

LangGraph (LANGGRAPH, 2024), an extension of LangChain, introduces a graph-based architecture for orchestrating multi-agent, stateful LLM applications. Unlike the linear pipelines in LangChain, LangGraph allows developers to define directed graphs where nodes represent LLM invocations or functions, and edges encode information flow. This structure supports cyclic processes and decision points, enabling agents to plan, execute, and reflect iteratively—an essential pattern for long reasoning chains and complex multi-step tool use. By maintaining application state and memory across cycles, LangGraph empowers developers to construct autonomous agents that move beyond one-shot prompting, thus advancing research in agentic AI.

2.8.2 N8N

N8N (ZAHRT, 2023) is an open-source workflow automation platform that enables users to connect diverse services and APIs into unified workflows with minimal coding. Through its visual drag-and-drop interface, users can design flows by linking nodes that represent specific actions—such as sending HTTP requests, transforming data, or interacting with cloud and database services.

Each node encapsulates a discrete task, and by chaining nodes together, users can orchestrate complex data pipelines and automation processes. In AI-driven contexts, N8N serves as a versatile orchestration layer, coordinating data collection, model inference, and result delivery within a single workflow.

For instance, an AI pipeline in N8N might include nodes for data input, preprocessing, model inference through external APIs, and post-processing or visualization of results. Decision-making nodes allow conditional routing of data, enabling the creation of adaptive, intelligent agents capable of responding to different scenarios dynamically.

By automating data flows and integrating multiple systems, N8N reduces manual intervention and accelerates deployment. Its monitoring tools support real-time supervision and performance tuning, making it suitable for both developers and non-technical users who aim to design efficient, maintainable AI workflows.

2.8.3 Retrieval-Augmented Generation

Retrieval-Augmented Generation (LEWIS et al., 2020a) is an advanced technique designed to enhance the accuracy and reliability of answers produced by large language models. Instead of relying exclusively on static information learned during the initial training phase, a RAG-based system dynamically searches for relevant information in external databases at the moment a question is asked. This design makes it possible to provide responses that are more precise, up-to-date, and grounded in verifiable data.

The process works in several steps, described below in simple terms:

1. **Generating representations from real-world data**: Initially, important documents (such as technical manuals, scientific articles, and expert reports) are collected and divided into smaller sections or text fragments. Each fragment is then converted into a numerical representation known as an *embedding*. This embedding serves as a mathematical summary of the content's meaning.

2. **Storing embeddings in a database**: The generated embeddings, along with links to their original text fragments, are stored in a specialized database called a vector database. This database allows the system to efficiently compare and retrieve information based on meaning rather than exact word matches.

3. **Encoding the user's query**: When a user submits a question, it is also transformed into an embedding using the same method as used for the stored documents. This transformation ensures that both the question and the stored content can be compared within the same semantic space.

4. **Performing similarity-based retrieval**: The system searches for embeddings stored in the database that are most similar to the query embedding. In practice, this involves identifying text fragments whose content is most relevant to answering the question. The search is based on mathematical similarity measures, such as cosine similarity.

5. **Recovering the original text**: Once the most relevant embeddings are identified, the system retrieves the corresponding original text fragments. These fragments contain the explicit information needed to construct a precise and evidence-based response.

6. **Enriching the prompt for generation**: Finally, the retrieved text fragments are included as additional context in the prompt provided to the language model. By incorporating this external knowledge, the model can generate responses that are not only contextually appropriate but also grounded in verified information, reducing the likelihood of errors or "hallucinations."

This method is particularly valuable in domains such as agronomy, where the reliability and correctness of recommendations are crucial. By integrating real-world data dynamically, RAG enables the development of AI systems that are more transparent, trustworthy, and aligned with scientific evidence.

2.8.4 Agents and Natural Language Interfaces

A major advancement in the field of generative AI is the development of agents that can interact with both structured and unstructured data using natural language (MOHAMMADJAFARI et al., 2024; TEAM, 2023; RICHARDS, 2023).

For structured data, these agents may use the data as direct input, or use a technique called text-to-SQL, which allows them to convert questions written in everyday language into SQL commands that can be run directly on databases. This makes it possible for people without technical training to explore and analyze data without needing to write code.

When dealing with unstructured data, such as documents, articles, or web pages, these agents combine retrieval techniques (like RAG) with reasoning abilities. They can search for relevant information, summarize it, and present clear answers to user questions.

This new way of interacting with data enables domain experts — for example, agronomists or plant pathologists — to access and analyze information directly, without always needing support from IT or data teams. Recent studies and user experiences show that natural language interfaces not only make data more accessible but also significantly improve productivity and support faster, more informed decision-making.

2.9 NATURAL LANGUAGE AS A UNIVERSAL INTERFACE FOR DATA

One of the most transformative developments in recent years is the adoption of natural language as a universal interface for interacting with both structured and unstructured data. Rather than relying on SQL queries, regular expressions, or manual pipeline engineering, users can now engage with databases, spreadsheets, documents, and APIs simply by describing their intentions in natural language. LLMs handle the translation into executable code, the execution itself, and the summarization of results. This paradigm significantly improves accessibility, empowering domain experts to work directly with data without requiring technical intermediaries.

This shift has fundamentally redefined the relationship between humans and data. Research workflows, business intelligence processes, and everyday automation now benefit from LLM-powered agents capable of executing commands, generating visualizations, and retrieving or synthesizing information—all through intuitive natural language interactions. As a result, the barrier to data-driven decision-making is lowered, enabling faster insights and more inclusive participation in data analysis tasks.

2.10 EMERGING TRENDS

As generative AI evolves from isolated models into collaborative, tool-using agents, new communication protocols have emerged to facilitate interoperability, negotiation, and secure execution across systems. Since late 2023, major organizations have introduced protocols designed to standardize how AI agents communicate and operate within multi-agent environments.

- **MCP (Model Context Protocol)** (ANTHROPIC, 2024): Released by Anthropic in late 2024, this client-server protocol is designed for model-to-tool invocation. It supports both stateless and session-aware interactions, using HTTP, Stdio, or Server-Sent Events (SSE) as transport layers. MCP is ideal for direct tool calls from language models.

- **A2A (Agent-to-Agent Protocol)** (GOOGLE, 2025): Introduced by Google in early 2025, A2A is a peer-to-peer protocol that enables agents to discover and negotiate with each other through HTTP-based “Agent Cards.” It assumes a centralized

agent directory and supports both stateless and session-aware communication, making it optimized for inter-agent cooperation.

- **AGP (Agent Gateway Protocol)** (CISCO, 2025): Developed by Cisco in 2024, AGP employs a gateway-based transport model with encrypted sessions. Utilizing gRPC (HTTP/2 + Protobuf), it provides secure routing between agents in high-throughput environments, making it well-suited for enterprise-scale deployments.

- **ACP (Agent Communication Protocol)** (ACP, 2025): Introduced by IBM in mid-2024, ACP adopts a brokered client-server architecture with registry-based discovery. It emphasizes tool modularity, session state tracking, and robust communication via HTTP streams, making it particularly suitable for distributed multi-agent architectures within organizations.

These protocols reflect an important shift in the generative AI ecosystem—from monolithic assistants to interoperable, modular agents capable of dynamic collaboration. As of 2025, they are still evolving and being tested in research and enterprise environments, but they lay the foundation for future standards in agent communication.

In parallel, several technological trends continue to shape the future of generative AI:

- **Tiny Models**: Lightweight models such as Phi-2, Mistral 7B, and TinyLlama offer strong performance for local and edge inference. Projects like Llama.cpp and GPT4All, launched in 2023, demonstrated that 7B–13B parameter models could run efficiently on laptops and mobile devices by leveraging 4-bit quantization and optimized inference libraries. This trend toward low-resource inference supports offline use cases and enables deployment in healthcare devices, vehicles, and other edge environments where connectivity may be limited.

- **Multimodal Interfaces**: Generative AI is increasingly extending beyond text to include vision, audio, and other modalities. Multimodal LLMs can now accept diverse inputs and produce outputs in multiple formats. GPT-4's vision extension, for example, allows the interpretation of images and diagrams, generating textual analyses or descriptions. Other systems, such as Bard and Bing Chat,

have integrated image understanding and generation (via models like DALL-E). Research models including BLIP-2, Flamingo, and PaLI have demonstrated approaches to connect vision encoders with LLMs, enabling tasks like visual question answering and image captioning. Speech capabilities have also advanced significantly: OpenAI's Whisper (speech-to-text) and new text-to-speech models have enabled voice-based interactions. By late 2023, ChatGPT supported voice dialogues, allowing users to converse naturally through spoken queries and synthesized responses. Google's Gemini is explicitly designed as a multimodal foundation model, capable of handling text, images, audio, and video within a unified framework.

- **Autonomous Agents**: Building on frameworks like LangChain and LangGraph, developers have started creating agents that can autonomously plan and execute multi-step workflows without continuous human prompting. A notable example is AutoGPT (2023), which chains GPT-4 instances to recursively break down goals into sub-tasks and solve them sequentially. Open Interpreter (2023) enables LLMs to execute code locally on a user's machine in response to natural language instructions, effectively allowing users to "talk to their computer." With appropriate safety measures and sandboxing, these autonomous agents hint at a future where personal AI assistants can perform complex, open-ended tasks on-device or online with minimal supervision.

- **Efficient Deployment**: As model capabilities grow, there is a parallel push to make them smaller, faster, and more accessible. Quantization techniques reduce the precision of model weights (and sometimes activations) from 16-bit to 8-bit, 4-bit, or even 3-bit integers, dramatically lowering resource requirements. Techniques like GPTQ (FRANTAR et al., 2023) allow large models (e.g., a 175B-parameter transformer) to be compressed post-training with minimal impact on accuracy, enabling deployment on a single GPU or even a high-end personal computer. Additional innovations such as sparsity (pruning redundant weights) and knowledge distillation (transferring knowledge from larger "teacher" models to smaller "student" models) further democratize access to generative AI by reducing operational costs and hardware barriers.

- **Offline and On-Device AI**: Tools like Llama.cpp and GPT4All make it feasible to run quantized LLMs on personal devices, unlocking use cases such as real-time AI translation on smartphones or vision-based processing in wearable devices like smart glasses.
- **Evaluation Infrastructure**: Frameworks like Ragas and LangSmith provide robust tools for evaluating and debugging LLM applications, supporting reproducibility, error tracking, and systematic performance analysis. This infrastructure is critical as AI systems become more autonomous and are integrated into mission-critical applications.

2.11 ADAPTATION OF THIS PROJECT

This work strategically integrates several of these recent advancements. Initial prototypes based on static prompts and rigid workflows were progressively replaced by LangChain pipelines, semantic search via pgvector, and LangGraph orchestration. The outcome is a flexible, modular agent capable of retrieving and reasoning over context in natural language, thereby delivering interactive, data-grounded responses that closely mimics expert consultations.

Moreover, by incorporating text-to-SQL and document question-answering agents, this project enables users to interact seamlessly with both structured data (e.g., PostgreSQL) and unstructured information (e.g., PDFs, images) through a unified conversational interface. The integration of multimodal capabilities and open-source models further enhances scalability and accessibility, ensuring the system remains robust and adaptable to diverse use cases.

2.12 CONCLUSION

Over the past five years, the role of AI in facilitating information access has been fundamentally redefined. By combining generative capabilities with retrieval, code execution, reasoning, and multimodal interaction, large language models now enable fluid and powerful engagement with both structured and unstructured knowledge sources. This chapter has outlined the key technological milestones that underpinned these ad-

vances and detailed how they have shaped the foundation of this project. As LLM tools continue to improve in performance, usability, and accessibility, the barriers between humans and data continue to dissolve, ushering in a new era of intelligent, conversational systems.

3 RELATED WORK

3.1 LARGE LANGUAGE MODELS IN AGRICULTURE

LLMs are increasingly recognized for their transformative potential across various sectors, and agriculture is no exception. More broadly, AI has seen rapid adoption in agriculture, supporting a wide range of applications such as automation, soil and crop monitoring, decision support, and resource optimization. AI-based solutions assist farmers in selecting optimal planting times, choosing suitable seeds for specific climate conditions, recommending soil nutrients, forecasting weather, and monitoring crop health in real time (ZHANG et al., 2021). These technologies contribute to increased productivity, reduced resource usage, and mitigation of environmental impacts (ZHANG et al., 2021).

The integration of LLMs in agricultural practices aims to enhance efficiency, improve decision-making, and address critical challenges such as crop monitoring and disease management. Automation in agriculture plays a vital role in these efforts, particularly through the development of early detection systems that can significantly reduce crop losses (ROUMELIOTIS et al., 2025). When combined with other advanced AI techniques, LLMs offer avenues to enhance the scalability and intelligence of precision agriculture systems, moving towards more automated and data-driven farming practices (ROUMELIOTIS et al., 2025).

3.1.1 Current Research on the Use of LLMs in Agriculture

Current research on the application of LLMs in agriculture spans a diverse range of tasks, from general crop monitoring to highly specialized areas like seed science. In parallel, various AI-powered technologies—such as sensors, drones, hyperspectral imaging, and agricultural robots—are being deployed to collect precise data on soil, climate, and crop health. These tools enable targeted interventions and automation of tasks like irrigation, spraying, and harvesting (WALEED et al., 2020; KUMAR et al., 2020). Intelligent monitoring systems provide farmers with detailed insights and tailored recommendations to maximize yield and optimize resource use (LIU, 2020).

A notable area of investigation involves the use of multimodal LLMs for automated

plant disease classification. For instance, studies have explored combining multimodal LLMs, specifically GPT-4o, with CNNs to detect plant diseases using leaf imagery (ROUMELIOTIS et al., 2025). This hybrid approach leverages the strengths of both model types: the LLM's understanding and generation capabilities with the CNN's proficiency in image analysis. Results from these investigations indicate that fine-tuned GPT-4o models can achieve performance comparable to, or even slightly better than, traditional deep learning models like ResNet-50. For example, fine-tuned GPT-4o models demonstrated up to 98.12% classification accuracy on apple leaf images, surpassing ResNet-50's 96.88% (ROUMELIOTIS et al., 2025).

This means that, in practice, farmers can expect more accurate and earlier detection of plant diseases using such systems, potentially preventing losses and reducing pesticide usage. Additionally, this integration shows promise for improved generalization and near-zero training loss, which can reduce the reliance on extensive labeled datasets and high-resolution sensor infrastructure, making advanced disease detection more accessible (ROUMELIOTIS et al., 2025).

Despite these advancements, the application of LLMs in highly specialized agricultural domains, such as seed science, remains nascent. This limitation is largely attributed to the scarcity of digital resources, the inherent complexity of gene-trait relationships, and a notable absence of standardized benchmarks for evaluating LLM performance in these niche areas (YING et al., 2025). More broadly, the adoption of AI in agriculture still faces challenges related to data quality, integration of new technologies into field operations, and the need to upskill farmers for effective use of these tools (AWASTHI, 2020; BELOEV et al., 2021). Even so, AI is considered essential for addressing rising food demand, labor shortages, and the impacts of climate change (CHEN et al., 2023).

To address these critical gaps, domain-specific benchmarks are being developed. SeedBench, for example, represents the first multi-task benchmark specifically designed for seed science. Developed in collaboration with domain experts, SeedBench aims to simulate key aspects of modern breeding processes, providing a structured environment for evaluating LLMs (YING et al., 2025). Benchmarks like SeedBench not only measure technical performance but also help ensure that these models are robust and reliable under realistic agricultural conditions. A comprehensive evaluation of 26 leading LLMs on SeedBench has revealed substantial discrepancies between the cur-

rent capabilities of general LLMs and the intricate demands of real-world seed science problems (YING et al., 2025). This initiative marks a foundational step, guiding future research and practical deployments in specialized agricultural domains.

3.1.2 NLP Applications in Agricultural Data Processing and Analysis

NLP applications are crucial for extracting actionable insights from the vast amounts of unstructured text data prevalent in agriculture, ranging from research papers and weather reports to farmer notes and market analyses. New frontiers in agricultural NLP involve investigating the effectiveness of pretraining transformer-based language models with extensive food-related text corpora (REZAYI et al., 2025). A notable example is AgriBERT, a domain-specific, fine-tuned, open-source model. AgriBERT has been trained from scratch using a large corpus of agricultural academic journals, comprising over 300 million tokens, to enable it to learn meaningful sentence representations specifically tailored for agricultural NLP applications (REZAYI et al., 2025). This approach addresses a key limitation of generic BERT models, which, when pretrained on general corpora like Wikipedia, may not generalize effectively across specialized domains due to their distinct vocabularies and contexts (REZAYI et al., 2023).

A significant application within agricultural NLP is semantic matching, which involves establishing accurate mappings between food descriptions and nutrition data (REZAYI et al., 2025). This task is critical for integrating diverse datasets, such as the USDA's Food and Nutrient Database with retail scanner data, to understand consumption patterns and inform public health policies (REZAYI et al., 2025). Fine-tuning domain-specific models like AgriBERT with external knowledge sources, such as the FoodOn ontology, enhances their ability to perform such semantic matching tasks (REZAYI et al., 2025). An exploratory investigation comparing AgriBERT with state-of-the-art general-purpose LLMs, including GPT-4, Mistral-large, Claude 3 Sonnet, and Gemini 1.0 Ultra, indicates that domain-specific models can effectively complement the broad knowledge and generative capabilities of these advanced LLMs in addressing the unique challenges of the agricultural sector (REZAYI et al., 2025). The integration of GPT-based models, either as a baseline for comparison or as an external knowledge source, further enhances AgriBERT's performance in semantic matching and its understanding of food-related concepts and relationships (REZAYI et al., 2025).

LLMs are also being explored for their utility in practical agricultural decision-making, particularly in pest management and the generation of diagnostic reports. Studies demonstrate that LLM-driven pest management decisions can achieve up to 72% accuracy when guided by instruction-based prompting that incorporates domain-specific knowledge (QIN et al., 2025). Furthermore, general LLMs like ChatGPT exhibit professional competence in analyzing agricultural data to generate accurate and timely reports, alerts, and insights, thereby facilitating informed decision-making and enhancing customer service within the agricultural domain (QIN et al., 2025). However, it is crucial to note that the accuracy of these predictions is heavily dependent on the quality of the input data. AI systems in agriculture are intended to assist decision-making and are not a substitute for human intuition and experience, especially in complex and dynamic agricultural environments (QIN et al., 2025). The YOLO-PC model, a lightweight variant of YOLO, further supports this by evaluating reasoning accuracy at 90% for agricultural diagnostic reports, emphasizing the importance of model-generated text quality in correlation with recognized information (QIN et al., 2025).

The synergy between domain-specific LLMs and general-purpose LLMs for optimal performance in agricultural NLP is a significant observation. The development of AgriBERT and its integration with advanced general LLMs like GPT-4 highlights a collaborative approach rather than a competitive one. AgriBERT provides a deep, nuanced understanding of agricultural terminology and concepts, while general LLMs offer broad knowledge and robust generative capabilities. This suggests that the future of specialized NLP applications, particularly in fields with unique vocabularies and contexts such as agriculture, will likely involve a layered architecture. This could entail fine-tuning smaller, domain-specific models on proprietary data for specialized tasks, while leveraging larger, general LLMs for broader reasoning, summarization, or user interaction, potentially through Retrieval-Augmented Generation (RAG) techniques. This also carries important cost implications, as training massive models from scratch for every niche domain is often impractical.

A critical consideration for LLM deployment in agriculture is the persistent importance of data quality and human oversight. While LLMs demonstrate promising accuracy rates, for instance, 72% in pest management, the explicit dependence of this accuracy on input data quality and the caveat that AI systems are not a substitute for human intuition and experience are crucial (QIN et al., 2025). This underscores that even

with advanced models, the principle of "garbage in, garbage out" applies, and human expertise remains indispensable for complex, real-world decision-making in dynamic agricultural settings. This observation implies that the successful deployment of LLMs in agriculture will necessitate robust data governance strategies, continuous human validation loops, and clear guidelines defining where AI serves as an assistant versus where human decision-making is paramount. It also highlights the need for user interfaces that facilitate the easy input of high-quality data and mechanisms for agricultural experts to review and, if necessary, override AI recommendations, especially in high-stakes scenarios like crop disease or pest management.

3.1.3 Connecting to This Dissertation

Building on these recent advances, this dissertation proposes a conversational plant diagnostic system that integrates the reasoning capabilities of large language models with robust image analysis. By leveraging multimodal approaches—combining textual symptom descriptions and leaf imagery—the system aims to provide accessible, real-time support to farmers and agronomists. Inspired by state-of-the-art studies, such as GPT-4o hybrid models and domain-specific benchmarks like SeedBench, this work addresses the urgent need for scalable, interpretable, and user-friendly AI tools in agriculture. The focus on transparency and adaptability seeks to bridge the gap between cutting-edge AI research and practical field applications, empowering small and medium-scale agricultural producers.

3.2 RAG FOR INFORMATION RETRIEVAL

Retrieval-Augmented Generation (RAG) has emerged as a prominent methodology, significantly enhancing the capabilities of LLMs by integrating dynamic information retrieval mechanisms into the generation process. This paradigm addresses key limitations of traditional LLMs, particularly their tendency to hallucinate or provide outdated information due to their static training data. RAG's ability to ground generative models in external, up-to-date knowledge sources has made it a focal point in natural language understanding and generation research.

3.2.1 Key Works on RAG and Its Application in Domain-Specific Knowledge Retrieval

Retrieval-Augmented Generation (RAG) has rapidly gained traction as a method to enhance the factual accuracy and adaptability of generative language models, especially when applied to specialized knowledge domains. By augmenting generative models with an external retrieval component, RAG systems can access and incorporate up-to-date or domain-specific information at inference time, reducing reliance on static pre-trained parameters and mitigating hallucinations (LEWIS et al., 2020b; SHI et al., 2023).

Lewis et al. (LEWIS et al., 2020b) introduced one of the foundational RAG frameworks, combining a dense passage retriever with a sequence-to-sequence generative model. This approach significantly improved performance on open-domain question answering benchmarks by allowing the model to dynamically incorporate retrieved evidence into its generated responses.

Guu et al. (GUU et al., 2020) proposed REALM, which integrates retrieval directly into pre-training, enabling the model to learn from external documents rather than only from its internal parameters. This architecture demonstrated strong improvements in both retrieval and generation accuracy, highlighting the benefits of retrieval-enhanced training for knowledge-intensive tasks.

In the context of domain-specific applications, recent studies have explored specialized retrieval corpora tailored to particular fields, such as medical guidelines, legal documents, or agricultural extension manuals (SHI et al., 2023; SUN et al., 2023). These works show that domain-adapted retrieval bases help address challenges like terminology ambiguity, specialized jargon, and the need for highly precise information.

Izacard and Grave (IZACARD; GRAVE, 2020) presented Fusion-in-Decoder (FiD), which extends RAG by fusing multiple retrieved passages within the decoder, allowing for richer context aggregation. FiD has been especially effective in tasks requiring synthesis of information from multiple sources.

In agriculture and plant health, RAG has the potential to provide real-time, evidence-backed responses to complex field queries. For example, a conversational assistant can retrieve the latest pest control guidelines or region-specific soil treatment protocols, offering practical support beyond what a purely parametric model can deliver.

Despite these advancements, challenges remain, including ensuring retrieval precision, integrating multimodal data (e.g., images and text), and optimizing latency for real-time deployments (SUN et al., 2023). Addressing these will be essential for the successful adoption of RAG-based systems in specialized fields like agriculture.

3.2.2 Comparative Analysis of RAG-Based Systems vs. Traditional Methods

Traditional information retrieval (IR) systems, such as keyword-based search engines and rule-based frameworks, rely heavily on exact lexical matches and static ranking algorithms. While efficient and interpretable, they often struggle to handle complex or nuanced queries, especially in specialized domains that involve evolving terminology and context-specific knowledge (VOORHEES; TICE, 1999).

Generative language models, on the other hand, can produce fluent and contextually rich responses but depend solely on internal parameters learned during pretraining. This reliance can lead to hallucinations and factual inaccuracies in knowledge-intensive tasks (JI et al., 2023). This issue is especially critical in fields like agriculture, where incorrect guidance can result in economic or environmental damage.

Retrieval-Augmented Generation (RAG) systems combine the strengths of both approaches by integrating external retrieval mechanisms with generative models. RAG systems retrieve relevant information dynamically at inference time, grounding responses in explicit evidence (LEWIS et al., 2020b). This enables them to produce more accurate, up-to-date, and context-sensitive answers while maintaining natural language fluency.

Table 1 summarizes key differences among traditional IR, generative models, and RAG systems.

Frame 1 – Comparison of Traditional IR, Generative Models, and RAG-Based Systems

Aspect	**Traditional IR**	**Generative Models**	**RAG-Based Systems**
Knowledge Source	Static indexed documents	Internal model weights	Retrieved external documents + model weights
Factual Accuracy	High (document-based)	Lower (prone to hallucination)	High (evidence-based grounding)
Fluency	Low (user interprets documents)	High	High
Explainability	High (explicit sources)	Low	High (retrieved evidence shown)
Adaptability to New Knowledge	Requires manual index update	Requires model re-training	Dynamic retrieval allows instant updates
Handling Complex Queries	Limited to keyword matching	Good contextual handling but risk of error	Strong contextual handling with factual evidence

However, RAG systems face challenges such as retrieval latency, dependence on retriever quality, and maintaining accurate, up-to-date external knowledge bases (SUN et al., 2023). Addressing these challenges is vital for their effective deployment in real-world agricultural contexts.

3.3 DEEP LEARNING MODELS FOR DISEASE CLASSIFICATION

The use of deep learning models has significantly advanced plant disease detection and classification, offering superior accuracy and scalability compared to traditional image processing and manual inspection methods. CNNs have become the backbone of image-based disease classification due to their ability to automatically learn hierarchical features from raw pixel data (SLADOJEVIC et al., 2016; MOHANTY; HUGHES; SALATHé, 2016).

Early works, such as those by Sladojevic et al. (SLADOJEVIC et al., 2016), demonstrated the feasibility of using CNNs to classify multiple plant diseases with high accuracy, even under varying environmental conditions. Mohanty et al. (MOHANTY; HUGHES; SALATHé, 2016) further extended this approach by training CNNs on a large dataset

containing images of healthy and diseased leaves across several crop species, achieving overall classification accuracies above 99% in controlled datasets.

Recent research has explored hybrid architectures that integrate CNNs with other deep learning modules to improve feature extraction and robustness. For example, combining CNNs with attention mechanisms or graph neural networks (GNNs) allows models to focus on critical lesion areas or to understand spatial relationships among disease patterns (LI et al., 2021; CHEN; LIU; WANG, 2023). Such hybrid models have shown enhanced performance in real-world field images, where occlusion, lighting variation, and background noise are common challenges.

Despite these advances, CNN-based approaches often require large, curated datasets for training, and their performance can degrade when deployed in diverse field conditions not represented in the training data (FERENTINOS, 2018). This motivates the integration of additional data modalities and adaptive mechanisms, such as transfer learning and multimodal frameworks, to enhance generalizability and reduce dependency on large annotated datasets.

In agricultural disease management, deep learning models facilitate rapid, large-scale monitoring and enable early intervention strategies, reducing yield losses and minimizing chemical input. However, practical deployment still requires models to be interpretable and adaptable to various crops and regions.

3.3.1 Studies on Deep Learning Models in Plant Disease Classification

Deep learning has revolutionized plant disease detection by enabling automatic feature extraction and robust classification from leaf images, outperforming traditional manual or rule-based methods (SLADOJEVIC et al., 2016).

3.3.1.1 Convolutional Neural Networks

Pioneering research by Sladojevic et al. (SLADOJEVIC et al., 2016) demonstrated that CNNs can achieve over 90% accuracy in multi-class plant disease identification. Mohanty et al. (MOHANTY; HUGHES; SALATHé, 2016) further validated this by training CNNs on a large dataset of healthy and diseased leaves, achieving above 99% accuracy in controlled settings.

A recent systematic review of over 160 studies from 2020–2024 highlights CNNs as the dominant architecture in plant disease detection, yet it notes challenges such as dataset diversity, model generalization, and deployment in natural environments (SUNIL; JAIDHAR; PATIL, 2023).

3.3.1.2 Attention-Enhanced Architectures

Attention mechanisms have been applied to highlight lesion regions, improving classification focus and robustness. For example, the APDC model leverages attention weighting to achieve up to 99.97% accuracy across multiple public datasets including PlantVillage and PaddyCrop (BERA; BHATTACHARJEE; KREJCAR, 2024a).

Vision Transformers (ViTs) have also begun to compete in this domain. The PlantXViT architecture—combining CNNs with ViTs—achieved 98–99% accuracy across crops like apple, maize, and rice, while also offering interpretability through visual attention maps (THAKUR et al., 2022).

3.3.1.3 Hybrid CNN–GNN and Graph-Based Models

To capture relational patterns, hybrid models combining CNNs with Graph Neural Networks (GNNs) have emerged. For instance, PND-Net integrates CNNs and GNNs for joint disease and nutrition deficiency classification, achieving 96–96.5% accuracy across multiple crops (BERA; BHATTACHARJEE; KREJCAR, 2024b). A soybean disease classification study combining MobileNetV2 with GraphSAGE reached 97.16% accuracy—surpassing single CNNs (95.04%)—while providing interpretable Grad-CAM visualizations (JAHIN et al., 2025).

3.3.1.4 Lightweight and Edge-Deployable Models

Efficiency-focused designs like Slender-CNN optimize parameter count and perform on par with heavier models (88–90% accuracy on corn, rice, wheat) while being suitable for deployment on resource-constrained devices (BAIJU et al., 2025).

MobileNetV3-based approaches achieved 99.66% classification accuracy on grape leaf diseases in real-time edge settings, demonstrating high precision (99.4%) and

viability for field use (PURANIK et al., 2024). Enhanced CNNs such as E-CNN also report 98% accuracy across crops like apple, corn, and potato (KUMARAN; SANJAY; SANTHIYA, 2024).

3.3.1.5 Transfer Learning, Data Augmentation, and Practical Challenges

Transfer learning has enabled broader model applicability across varied datasets, while targeted data augmentation and preprocessing significantly improve model robustness against lighting variations and environmental noise (PAWAR et al., 2024).

However, real-world deployment faces challenges: limited field data, the need for segmentation, and model interpretability. Issues like occlusion, small datasets, and lab-to-field transitions remain significant hurdles, as emphasized in recent literature (LIU; WANG, 2021; SUNIL; JAIDHAR; PATIL, 2023).

3.3.2 Studies on Deep Learning Models in Tomato Leaf Disease Classification

Several recent studies have benchmarked deep learning architectures on the PlantVillage tomato leaf dataset, reporting consistently high accuracy while highlighting trade-offs between model complexity and performance. Table 1 consolidates representative results from key publications:

Table 1 – Comparative Performance of Deep Learning Models for Plant Disease Classification on PlantVillage Tomato Leaf Dataset

Model	Accuracy (%)	Citation
ResNet50-DPA	97.60	(LIANG; JIANG, 2023)
MaxViT	97.00	(GHOSH et al., 2025)
EfficientNet Ensemble	96.99	(GONZALES; DIOSES, 2024)
Hybrid CNN–Transformer	95.22	(NEMMOUR et al., 2025)

These results reinforce several key observations:

- **CNNs remain strong performers:** ResNet-style architectures augmented with spatial attention (e.g., ResNet50-DPA) achieve near 98% accuracy under controlled conditions (LIANG; JIANG, 2023).

- **Vision Transformers show promise:** Transformer-based models like MaxViT approach CNN-level performance (97%) in classification tasks (GHOSH et al., 2025).
- **Hybrid and ensemble methods:** Combining CNN and transformer features yields high accuracy (96–97%) with more efficient resource use (GONZALES; DIOSES, 2024; NEMMOUR et al., 2025).

While PlantVillage enables strong benchmark performance, these models are typically trained and evaluated on clean, uniform-background images. Adapting them to field deployment remains a challenge due to occlusion, variability in lighting, and diverse backgrounds. This motivates the adoption of lightweight models and RAG-integration in this dissertation, to ensure robustness and explainability in real-world agricultural diagnostics.

3.3.3 Application of CNNs and Other Architectures in Agricultural Diagnosis

CNNs have become the cornerstone of automated plant disease detection, providing strong performance across diverse crop types and imaging conditions. Recent reviews report that CNN-based models—including ResNet, EfficientNet, and VGG variants—dominate the space, often achieving 90–99% accuracy on benchmark datasets such as PlantVillage (CHEN et al., 2024a).

To enhance both accuracy and interpretability, researchers have explored hybrid models combining CNNs with transformers or attention modules. For instance, a hybrid CNN–Transformer model achieved 99.45% accuracy on tomato leaf disease classification using CycleGAN-augmented data and attention-enhanced feature extracts (CHEN et al., 2024b). Similarly, models like CMTNet and FOTCA—integrating CNN and transformer modules—have excelled in fine-grained and robust field scenarios, with performance exceeding 99% accuracy (GUO; FENG; GUO, 2025; HU et al., 2023).

Table 2 highlights key architectures and their performance on plant disease datasets:

Table 2 – Performance of CNN and Hybrid Architectures in Plant Disease Detection

Model	Description	Accuracy (%)
Advanced CNN + SE blocks	CNNs with attention and residual enhancements	99.39
Hybrid CNN–Transformer (CycleGAN)	Combines CNN features with transformer context	99.45
CMTNet	CNN + Transformer for spectral-spatial classification in UAV images	>99
FOTCA	Transformer-CNN hybrid with Fourier-based attention	99.8

- **CNNs with attention enhancements** (GONZÁLEZ-BRIONES et al., 2025) balance high accuracy and interpretability, focusing on symptom regions in images.
- **CNN–Transformer hybrids** (CHEN et al., 2024b) benefit from global context and local feature fusion, improving robustness in challenging image conditions.
- **Transformer-heavy models** like CMTNet(GUO; FENG; GUO, 2025) and FOTCA (HU et al., 2023), while achieving top performance, require more complex integration and data preprocessing.

Despite excellent benchmarks, these models must be adapted for real-world use — handling noisy images, diverse environments, and minimal annotated data. The approach in this dissertation adopts an efficient CNN backbone with attention mechanisms, complemented by RAG to provide contextual, explainable recommendations — enhancing both robustness and practical utility in agricultural diagnostics.

3.4 USE OF AI ASSISTANTS IN DIAGNOSIS SYSTEMS

AI assistants—ranging from image-enabled chatbots to retrieval-augmented multimodal agents—are transforming diagnostic tools in both medicine and agriculture by combining visual analysis, conversational interaction, and curated knowledge retrieval.

In healthcare, AI assistants such as ChatGPT, OpenAI's HealthBench, and Microsoft's Diagnostic Orchestrator have demonstrated remarkable diagnostic capabilities. In one clinical evaluation, AI alone achieved a median diagnostic reasoning score of 92%, outperforming physicians (76%) and physician–AI teams (74%), although ac-

curacy dropped when physicians interacted with the AI (GOH et al., 2024). Another study reported Microsoft's Diagnostic Orchestrator achieving 85.5% accuracy on 304 medical cases, vastly outperforming a physician panel at 20% accuracy (SALLAM et al., 2025). These results illustrate high potential but also highlight risks such as hallucinations, automation bias, and reduced trust when AI disagrees with human experts (GOURABATHINA et al., 2025).

In agriculture, systems combining CNN-based leaf image analysis with conversational, knowledge-grounded interfaces have shown promise in supporting disease diagnosis and farm management. For example, YOLO-integrated RAG systems have enabled early, context-aware detection of diseases in coffee and medicinal plants (KUMAR et al., 2024). Recent reviews confirm that integrating image processing with retrieval-augmented reasoning improves diagnostic precision and interpretability (TU et al., 2025).

Key advantages of AI assistants include:

- **Improved accuracy:** Grounding responses in external evidence and multimodal inputs reduces hallucinations and enhances reliability (SINGHAL et al., 2023).
- **Enhanced user interaction:** Conversational interfaces allow for plain-language explanations and actionable guidance, making tools accessible to both farmers and clinicians (GOH et al., 2024).
- **Scalability and updatability:** Retrieval-based models can incorporate new information without full retraining, critical for rapidly evolving diseases and practices (TU et al., 2025).

Nevertheless, important challenges persist:

- **Automation bias:** Users may over-rely on AI suggestions without critical appraisal, potentially leading to misdiagnoses or mismanagement (GOH et al., 2024).
- **Hallucinations and factual errors:** AI systems can generate plausible but incorrect information, with some studies estimating hallucination rates between 27% and 47% (JI et al., 2023).
- **Equity and trust:** Bias in model training and a lack of transparency can hinder adoption in diverse agricultural contexts, especially where local knowledge and dialects are crucial (SINGHAL et al., 2023).

Integration in This Dissertation

The system proposed in this dissertation integrates a CNN-based image classifier with a Retrieval-Augmented Generation assistant. By grounding every diagnostic suggestion in visual evidence and explicit knowledge sources, the system delivers accurate, context-rich, and explainable recommendations. This approach emphasizes clarity and trust, empowering agronomists and farmers with a robust, field-ready decision-support tool.

3.5 PLANT DISEASE DIAGNOSIS: CONCEPTS AND PRACTICES

Accurate plant disease diagnosis is a cornerstone of effective crop protection and sustainable agriculture. It enables timely intervention, reducing yield loss and unnecessary pesticide use. The diagnostic process typically follows a structured flow: **observation** of symptoms, formulation of a **hypothesis** regarding potential causes, **confirmation** through targeted testing, and delivery of a **recommendation** for management (RILEY; WILLIAMSON; MALOY, 2002). Early and precise diagnosis is essential to minimize costs and prevent pathogen spread (ALI, 2022).

Conventional Methods

Traditional diagnosis relies on visual examination of **symptoms and signs**, which remains the most accessible approach for many practitioners (RILEY; WILLIAMSON; MALOY, 2002). Its main advantages are low cost and speed, particularly for diseases with distinctive visual markers. However, it is inherently subjective and depends heavily on expert experience. Many diseases present overlapping or nonspecific symptoms, complicating accurate identification.

Molecular and Serological Methods

Molecular techniques such as PCR, quantitative PCR (qPCR), and loop-mediated isothermal amplification (LAMP) offer high sensitivity and specificity for pathogen detection (GOMEZ-GUTIERREZ; GOODWIN, 2022; NÉMETH; KOVÁCS, 2025). While LAMP is

suitable for rapid, field-ready testing due to its low equipment requirements, it faces challenges related to primer design and contamination risk (NÉMETH; KOVÁCS, 2025). Serological methods like enzyme-linked immunosorbent assays (ELISA) and immunochromatographic strips (lateral flow tests) are widely used for detecting plant viruses, offering fast and moderately inexpensive results. Yet, their reliability depends on antibody quality and can be affected by cross-reactivity and false positives (KANAPIYA et al., 2024; KIM et al., 2024).

Sensor-Based and Imaging Approaches

Emerging approaches use **RGB and thermal sensors** to identify disease-induced stress non-destructively. RGB imaging enables the extraction of color and texture features from visible-light images, while thermal cameras capture canopy temperature anomalies associated with water stress or infection (WALSH; MANGINA; NEGRÃO, 2024). These techniques are increasingly integrated with drones and smartphones, making them scalable and accessible. However, their accuracy can be influenced by external factors such as lighting conditions, crop variety, and environmental variability (WALSH; MANGINA; NEGRÃO, 2024). When combined with deep learning or multimodal AI frameworks, these sensor-based systems are paving the way for large-scale, automated plant health monitoring.

Overall, plant disease diagnosis has evolved from qualitative observation toward **quantitative, multimodal, and AI-assisted** methodologies. These advancements are crucial for developing intelligent systems capable of delivering fast, evidence-based, and scalable diagnostics—aligning directly with the goals of this dissertation.

3.6 SUMMARY

LLMs have rapidly emerged as transformative tools across many industries, and agriculture is no exception. By leveraging advanced natural language understanding and generation capabilities, LLMs can assist in decision support, advisory services, and knowledge dissemination to farmers and agronomists.

Recent studies have demonstrated the utility of LLMs for a wide range of agricultural tasks, including personalized crop management advice, pest and disease di-

agnosis, climate adaptation strategies, and supply chain optimization (SAPKOTA et al., 2024; BREZULEANU et al., 2025). For example, fine-tuned versions of GPT models have been deployed to answer natural language questions from farmers, providing region-specific recommendations on irrigation, fertilizer use, and pest control (BREZULEANU et al., 2025).

One of the key advantages of LLMs in agriculture is their ability to democratize access to expert knowledge. In many rural regions, the scarcity of agronomists limits timely and accurate guidance. LLMs can help bridge this gap by providing farmers with instant, easy-to-understand responses in local languages or dialects, thereby empowering decision-making at the farm level (SAPKOTA et al., 2024).

Additionally, LLMs have been integrated with satellite data, IoT sensors, and image analysis pipelines to create holistic farm monitoring systems. Such systems can generate comprehensive reports that merge textual insights with real-time environmental and crop health data, enabling precision agriculture at scale (TENG et al., 2023; JINDAL; KAUR, 2024).

Despite these advancements, challenges remain. LLMs are susceptible to hallucinations—generating plausible but incorrect information—and require continuous grounding in up-to-date agronomic data to ensure reliability (JI et al., 2023). Moreover, integrating LLMs into field workflows demands careful consideration of user trust, explainability, and cultural acceptance, especially in diverse agricultural contexts (BREZULEANU et al., 2025).

Integration in This Dissertation

This dissertation builds upon these insights by employing a retrieval-augmented LLM framework, which combines plant image analysis with textual reasoning grounded in a curated agricultural knowledge base. This integrated approach aims to provide precise, explainable, and context-aware recommendations, directly addressing common limitations in both purely vision-based and purely language-based diagnostic tools.

3.7 FINAL CONSIDERATIONS

This chapter reviewed the evolution and application of advanced AI techniques — particularly deep learning models, hybrid architectures, and large language models — in the context of plant disease diagnosis and agricultural decision support. Early CNN-based approaches demonstrated strong performance in controlled image classification tasks, while recent hybrid models combining transformers, attention mechanisms, and graph networks have further improved robustness and interpretability.

The integration of LLMs and retrieval-augmented generation frameworks represents a significant leap forward in providing context-aware, evidence-backed, and scalable diagnostic assistance. These systems enable real-time, personalized support for farmers and agronomists, addressing gaps left by purely vision-based or rule-based methods. Moreover, they democratize access to expert-level guidance, particularly in regions with limited technical resources.

Despite these advances, critical challenges remain—including ensuring generalization under diverse field conditions, mitigating hallucinations, and fostering user trust through transparent and interpretable outputs. Addressing these challenges requires careful design, continuous grounding in domain knowledge, and a focus on explainability.

Building upon these insights, this dissertation proposes a novel multimodal system that combines CNN-based image analysis with RAG-enhanced conversational reasoning. By grounding visual diagnoses in an up-to-date agricultural knowledge base and delivering clear, context-rich explanations, the proposed approach aims to provide accurate, practical, and trustworthy decision support to agricultural stakeholders. This system aspires to bridge the gap between cutting-edge AI research and real-world farming needs, ultimately contributing to more resilient and sustainable agricultural practices.

4 METHODS

4.1 PROPOSED METHODOLOGY

4.1.1 Hybrid System: RAG and Deep Learning

The proposed system combines RAG and deep learning for images to support robust and accurate plant disease diagnostics. RAG grounds the system's responses in a curated agronomic knowledge base, significantly mitigating hallucinations and improving factual accuracy. Meanwhile, deep learning models provide visual diagnosis capabilities from plant images, acting as a complementary "eye" for the conversational assistant.

The knowledge base was constructed from authoritative agronomy books and technical manuals. Texts were digitized via optical character recognition (OCR), manually cleaned to remove artifacts, and segmented into chunks of approximately 1,000 characters with a 200-character overlap to preserve context. These chunks were then embedded using OpenAI's embedding models and stored in a PostgreSQL database extended with pgvector. During diagnosis, a retriever implemented with LangGraph and cosine similarity retrieves the most relevant knowledge snippets to inform and ground the responses.

4.1.2 Reasoning Techniques

To enhance reasoning capabilities, our system employs the Chain-of-Thought technique (ZHANG et al., 2024). This approach encourages the model to think step-by-step through problems, improving its ability to solve complex tasks by breaking them down into smaller, manageable steps.

4.1.3 Multi-modal Capabilities

Our system integrates multi-modal capabilities (XIE et al., 2024), allowing it to process and generate content across different media types, such as text, images, and audio. This versatility broadens the system's applicability and enhances its performance

in diverse tasks.

4.1.4 Framework Selection: LangChain, LangGraph

We briefly evaluated the leading agentic and RAG frameworks—including CrewAI, Phidata, and LangGraph—to inform our tooling choices. Our evaluation criteria focused on development speed, flexibility, community support, and ability to build complex workflows.

4.1.4.1 CrewAI

CrewAI (CREWAI, -) facilitates the creation of collaborative, role-based multi-agent systems with relatively low code overhead. However, its higher-level abstractions were too constrained for our need to tightly control query transformation, retrieval, reranking, and generation logic.

4.1.4.2 Phidata

Phidata (PHIDATA, -) offers built-in RAG capabilities, multi-modal support, and a clean API ideal for prototypes. However, it proved limiting when extending beyond basic use cases: building our specific diagnostic pipeline on top of its abstractions required extensive workarounds and reduced configurability for components like custom reranking and prompt templates.

4.1.4.3 LangChain & LangGraph

LangChain (LANGGRAPH, -) was selected as the foundation due to its broad ecosystem—including connectors, prompt templates, and vector store integrations—as well as its modular chaining paradigm. We then added LangGraph for orchestration because it provides graph-based workflow control, visual debugging, and explicit state management suitable for our multi-stage RAG pipeline. In contrast to CrewAI's abstraction, LangGraph gave us the fine-grained control we needed over each pipeline component. Additionally, its active open-source community ensures long-term viability

and continued feature development.

4.1.4.4 Summary

Overall, we prioritized frameworks that (1) enabled rapid prototyping, (2) supported modular customization at every stage, and (3) matched community maturity. LangChain + LangGraph best fit these criteria—whereas CrewAI felt restrictive and Phidata, while easy to start with, lacked the flexibility required for our full diagnostic system.

4.2 SYSTEM DESIGN AND ITERATIVE PROTOTYPING

The development of the system followed an iterative, prototype-driven methodology, shaped by rapid advancements in large language models and agent frameworks. These iterations were strategically designed to address the central research objective of minimizing hallucinations and increasing diagnostic accuracy in plant disease responses.

4.2.1 First Prototype

The first prototype, built with N8N, relied on a straightforward workflow that integrated a simple *semantic-only* RAG pipeline. While it allowed rapid experimentation, this approach often produced hallucinations and lacked the flexibility to incorporate new tools or improve retrieval strategies. Figure 1 shows the basic architecture.

Despite enabling fast deployment, the first prototype's single-layered retrieval limited both the depth and adaptability of responses, motivating further iterations.

The system is composed of multiple components connected through a workflow orchestrated in n8n. Below, we describe each component and its interaction in detail.

4.2.1.1 User Interaction

The entry point is the **Telegram Trigger** node. When a user sends a message (text, photo, or audio) to the bot on Telegram, this node captures the content and initiates the workflow.

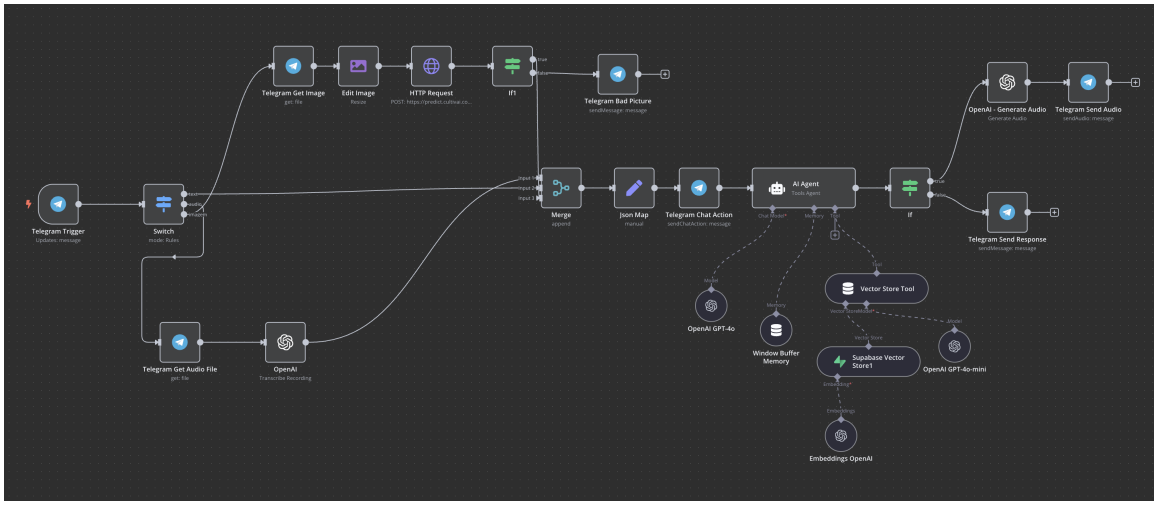

Figure 1 – First prototype architecture diagram implemented in n8n.

4.2.1.2 Input Type Determination

Next, the **Switch** node analyzes the message to determine its type: text, audio, or image.

- **Text**: Sent directly to be merged and processed later.
- **Audio**: The system fetches the audio file via the **Telegram Get Audio File** node and transcribes it using the **OpenAI Transcription** node.
- **Image**: The image is retrieved using the **Telegram Get Image** node and resized or pre-processed with the **Edit Image** node.

4.2.1.3 Image Analysis

If an image is provided, it is further analyzed by sending it to an external plant disease prediction API through the **HTTP Request** node. This external service returns predictions on whether the plant is sick, possible pathogens, symptoms, and the type of plant.

A condition (**If1** node) checks if the returned analysis contains a valid plant classification. If valid, the system proceeds; otherwise, it notifies the user that the image is not suitable.

4.2.1.4 Merging Input Data

The different possible inputs (text, audio transcription, image analysis results) are merged in the **Merge** node to form a unified content object.

4.2.1.5 Mapping Predictions

After merging, the **Json Map** node organizes predictions into structured fields, such as `is_sick`, `pathogen_predictor`, and `symptom_predictor`.

The system then sends an initial chat action signal to the user via the **Telegram Chat Action** node, indicating that the bot is preparing a response.

4.2.1.6 Agent Reasoning and Memory

The core reasoning is performed in the **AI Agent** node. This node uses a detailed prompt specifying that the agent should consult the internal vector store database (CultivAI knowledge base) and avoid generating speculative content. The memory context is maintained using the **Window Buffer Memory** node, which ensures the conversation continuity for each user session.

4.2.1.7 Knowledge Base and Embeddings

The system integrates a Supabase vector store (**Supabase Vector Store1**) and a **Vector Store Tool** node to store and query embeddings generated via **OpenAI Embeddings**. These components allow efficient similarity search against pre-stored expert knowledge, improving grounding and reducing hallucinations.

4.2.1.8 Language Model Integration

The **OpenAI GPT-4o** node is used to generate or refine responses, and is tightly integrated with the AI Agent node to maintain context and ensure high-quality outputs. A smaller model (**OpenAI GPT-4o-mini**) is connected to support auxiliary tasks in the vector store.

4.2.1.9 Final Response and Output Options

The final response is processed through a condition node (**If**). Depending on whether the output includes audio or text, the system either:

- Generates audio using the **OpenAI - Generate Audio** node and sends it back via the **Telegram Send Audio** node.
- Sends a textual reply through the **Telegram Send Response** node.

If an image is invalid, a fallback message is sent via the **Telegram Bad Picture** node.

4.2.2 Second Prototype: Multi-Agent Architecture

The second LIMMO prototype, developed in early 2025, introduced a more advanced architecture inspired by a team-of-specialists approach. Instead of relying on a single workflow, this version used multiple virtual agents, each with a defined area of expertise, working under the supervision of a main controller.

Figure 2 illustrates the conceptual architecture. When a user sends a query (text, image, or voice), the main controller (supervisor) analyzes the initial message and routes it to the appropriate team of virtual specialists.

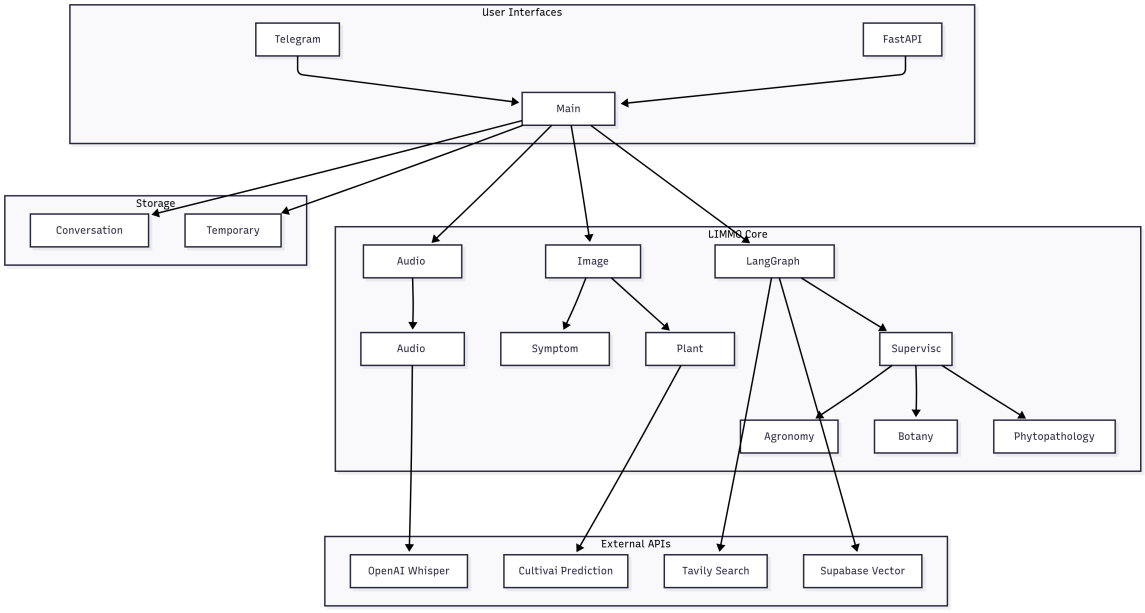

Figure 2 – Supervisor component architecture diagram.

The system included three main specialist teams:

- **Phytopathology Team**: Focused on diagnosing plant diseases, with sub-specialists such as a mycologist (fungal diseases), virologist, and bacteriologist.
- **Botany Team**: Responsible for plant identification and understanding plant-environment interactions.
- **Agronomy Team**: Provided advice on soil management, climate impacts, and cultivation techniques.

The image and audio inputs were processed by dedicated modules: an image analyzer (using prediction APIs and internal models) and an audio transcriber (using OpenAI Whisper API).

For text-based reasoning and knowledge integration, the system used a combination of a vector database (Supabase with embeddings) and web search APIs (e.g., Tavily). These tools allowed each team to access relevant agronomic literature and external data as needed, particularly when information was not available in the local knowledge base.

The multi-agent design was inspired by the metaphor of a real plant health clinic, where different specialists collaborate to give comprehensive support. While conceptually rich, this architecture became increasingly challenging to scale. The growing number of agents and complex prompt coordination led to qualitative difficulties during testing: maintenance burdens increased, prompt drift was frequent, and debugging became frustratingly slow. Although not formally quantified, these challenges aligned with known limitations reported in recent multi-agent AI frameworks, which cite maintenance overhead and coordination latency as key bottlenecks (SHI et al., 2023).

This experience provided valuable insights and led to the decision to adopt a more streamlined single-agent approach in the subsequent version.

4.2.3 Final Prototype: Single-Agent RAG-Enhanced Architecture

Learning from earlier iterations, the final design shifted toward a streamlined single-agent architecture augmented with RAG and modular tools (MCP). This simplified design was easier to maintain, debug, and extend. By consolidating responsibilities into

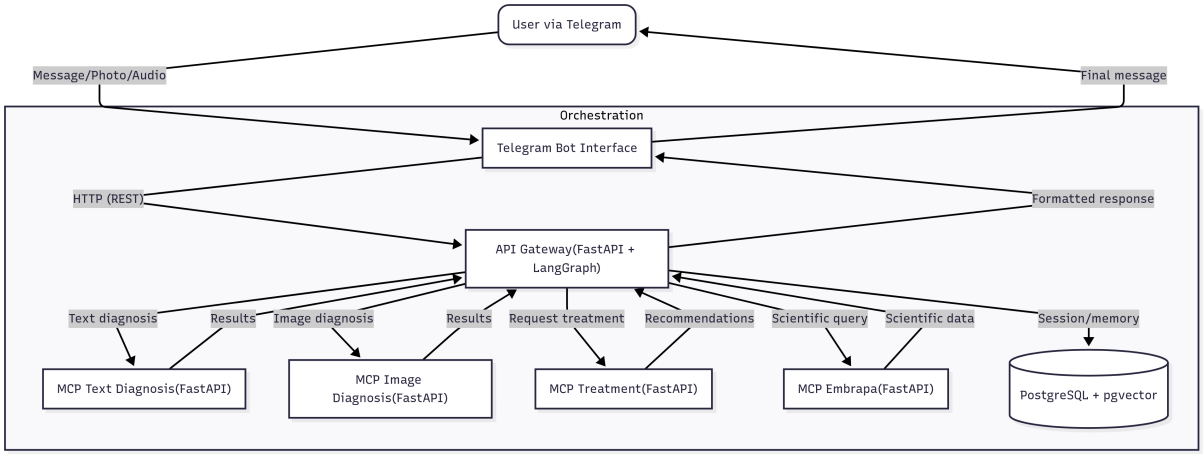

Figure 3 – Final architecture data flow diagram.

a single orchestrated agent, it became straightforward to monitor, update prompts, and integrate additional capabilities.

The final architecture also improved robustness and efficiency. Responsibilities were consolidated into a single orchestrated agent, making monitoring, debugging, and prompt updates much simpler. Moreover, the design leveraged modular tools implemented via Model Context Protocol (MCP), allowing seamless integration of external services such as image analysis APIs and future agronomic data sources. For example, images are uploaded to Google Cloud Storage and then passed as secure links to the agent, which uses an image analysis MCP to detect symptoms and probable diseases. Audio inputs are pre-processed before agent handling, maintaining modularity and extensibility.

The use of MCP facilitates future expansion, as new tools or data sources can be integrated with minimal changes to the core agent logic, aligning with best practices in scalable AI system design.

The CNN module provides a preliminary analysis of detected symptoms and candidate diseases. This output serves as guidance for the RAG module, which then retrieves and reasons over knowledge snippets to produce a grounded and context-rich response. Importantly, users can engage conversationally to confirm or clarify symptoms, thus improving overall diagnostic accuracy and user trust.

Performance improvements were qualitatively significant: hallucinations decreased markedly thanks to the hybrid retrieval strategy (combining semantic and keyword-enhanced approaches), while response times improved due to the reduction in agent orchestration overhead. These observations align with recent findings on improved factual consistency and latency reductions in simplified RAG-enhanced architectures (JI

et al., 2023; SUN et al., 2023; SHI et al., 2023).

4.3 AI FRAMEWORKS AND DESIGN DECISIONS

LangChain and LangGraph(LANGGRAPH, -) were chosen for their accessible learning curve, high-quality documentation, and modular design, which facilitated rapid prototyping and integration of complex workflows. Alternative frameworks such as Phidata and CrewAI were evaluated but ultimately set aside due to implementation and documentation limitations.

The adoption of RAG played a crucial role in grounding answers in the domain-specific knowledge base, significantly reducing hallucination rates and enhancing user trust—an essential requirement in agricultural diagnostics.

4.4 SYSTEM ARCHITECTURE OVERVIEW

The final architecture of LIMMO is a modular, containerized system providing AI-powered plant disease diagnosis through a Telegram bot interface. The architecture prioritizes scalability, reliability, and maintainability, using Python, FastAPI, and container orchestration via Docker Compose.

4.4.1 High-Level Architecture

The system consists of the following main components:

- **Telegram Interface**: Manages user interactions, including text, voice, and image messages, and formats responses.
- **AI Agent (LangChain MCP Agent)**: Central orchestrator for query routing, memory management, and reasoning workflows.
- **Diagnosis MCPs**: Services for text-based and image-based diagnosis, treatment recommendations, and external agricultural knowledge integration.
- **Database**: PostgreSQL with pgvector for embeddings and session memory.

- **Cloud Infrastructure**: Managed with Docker Compose to enable scalability and deployment flexibility.

4.5 COMPONENT IMPLEMENTATION DETAILS

4.5.1 Telegram Interface

Implemented in Python with `python-telegram-bot`, this interface handles incoming messages, converts audio to text using OpenAI Whisper API, manages sessions, and formats AI responses to comply with Telegram's presentation rules.

4.5.2 AI Agent

Built using Python, FastAPI, LangChain, and LangGraph, the AI agent serves as the core orchestrator. It manages context, orchestrates calls to MCPs, and integrates RAG pipelines to combine retrieved knowledge with language model reasoning.

The agentic retriever uses similarity-based searches (cosine similarity) on vector embeddings to ground responses effectively.

4.5.3 Evolution of RAG Approaches

Throughout the development of the system, three distinct RAG approaches were evaluated. Each approach shared the same underlying tools, including OpenAI embeddings, Facebook AI Similarity Search (FAISS)(JOHNSON et al., 2019) as the vector store, and LangChain components for orchestration. These iterations were designed to address the main research objective of reducing hallucination and improving factual accuracy in plant disease diagnosis responses.

4.5.3.1 Semantic-Only RAG

The first prototype employed a simple **semantic-only retrieval strategy**. In this approach, user queries and documents were both embedded using OpenAI embeddings and matched purely on semantic similarity within the FAISS index. While this allowed

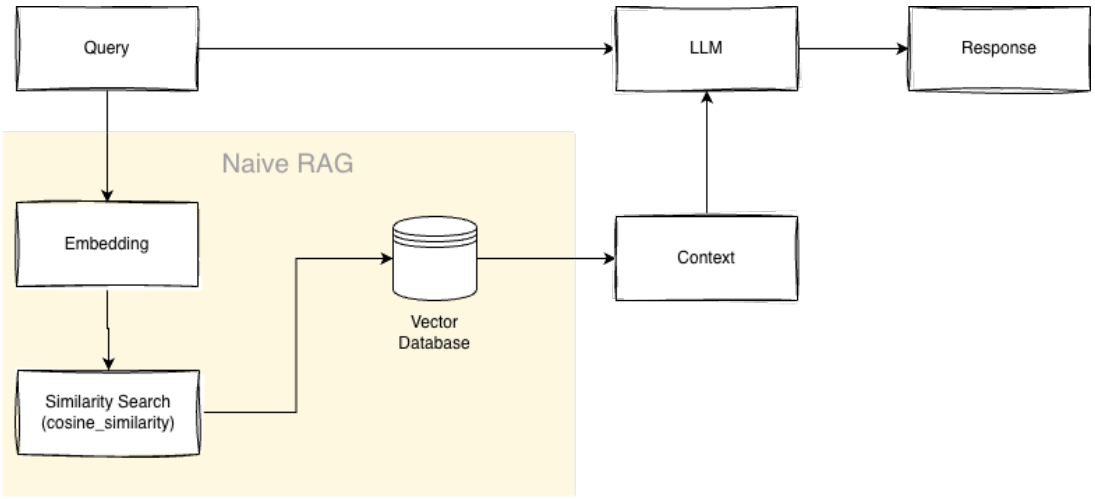

Figure 4 – Semantic-only RAG pipeline. The user query and documents are embedded and matched by cosine similarity in a vector index. The top-k chunks are passed directly to the LLM to compose the answer. This design is simple and fast, but can retrieve semantically close yet off-topic passages because it lacks keyword checks and reranking.

for a straightforward and fast implementation, it sometimes retrieved documents that were semantically close but not necessarily factually relevant or precise, especially in cases where the query contained ambiguous or broad terms. The pipeline is depicted in Fig. 4.

4.5.3.2 Hybrid Semantic and Keyword RAG

The second prototype introduced a **hybrid retrieval strategy** that combines semantic similarity with traditional keyword-based search. In this version, initial results from semantic retrieval were cross-verified using keyword matching heuristics. This combination helped to improve the precision of document selection, particularly when specific technical terms or disease names were present in the query. By integrating keyword filtering, the system reduced off-topic retrievals and improved factual alignment of the responses. The pipeline is depicted in Fig. 5.

4.5.4 Agentic RAG Implementation

To ensure accurate and contextually grounded responses, the system integrates a Retrieval-Augmented Generation approach. This architecture allows the language model to retrieve domain-specific documents and incorporate them into its reasoning process, mitigating hallucination, and improving factual consistency. The pipeline is

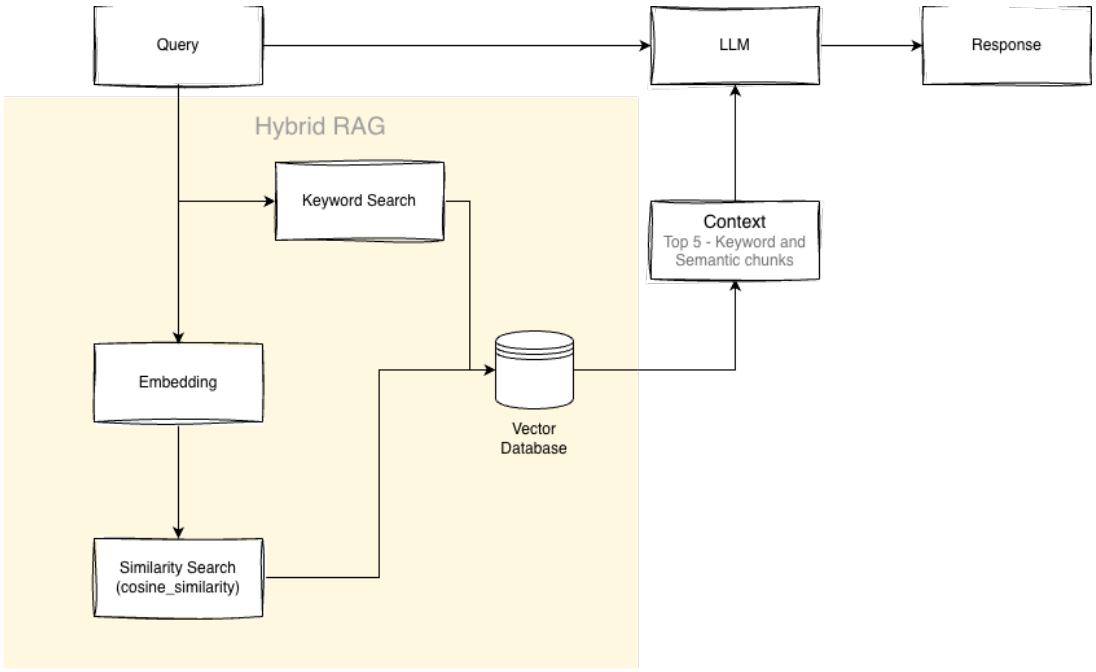

Figure 5 – Hybrid RAG pipeline. Semantic retrieval is complemented by keyword-based search using domain terms (crop, pathogen, symptom). The final context is formed from passages that are both semantically similar and keyword-confirmed, improving precision and reducing irrelevant chunks compared to semantic-only retrieval.

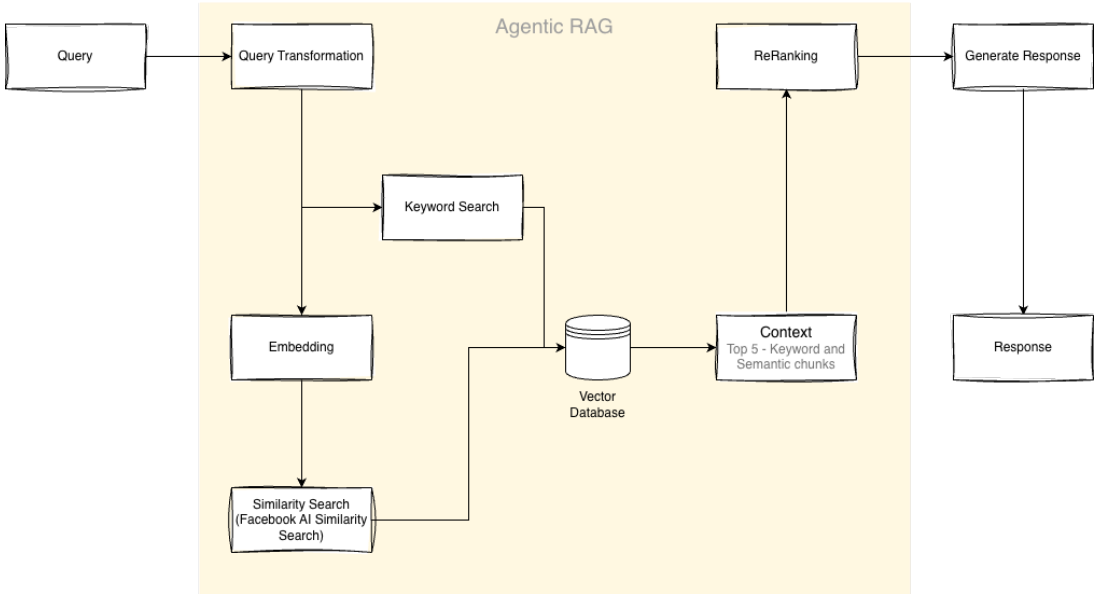

Figure 6 – Agentic RAG pipeline. A controller transforms the query, runs multi-path retrieval (semantic, keyword, and optional metadata filters), and applies reranking before building a structured context window with citations. Orchestrated with LangGraph, the LLM is instructed to ground answers in the retrieved evidence and to state uncertainty when information is insufficient, reducing hallucinations while maintaining clarity.

depicted in Fig. 6.

4.5.4.1 Embedding Model for Retrieval

The system uses **OpenAI Embeddings** to convert textual documents and user queries into vector representations. This is implemented through the `OpenAIEmbeddings` class from the `langchain_openai` package, using the OpenAI model text-embedding-3-small with 1536 dimensions, initialized in the `AgenticRetriever` class with default parameters. These embeddings provide a semantically rich representation of plant-related content, enabling effective similarity-based search.

4.5.4.2 Indexing Strategy

For indexing and similarity search, the system employs **Facebook AI Similarity Search** as the vector store backend. Documents are indexed using the `FAISS.from_documents()` method in combination with OpenAI embeddings. The retrieval process is configured to use a top-k similarity search, with k set to 5 by default. This configuration ensures that the most relevant five documents are retrieved for each query.

Additionally, the system supports incremental updates to the knowledge base via the `add_documents()` method, enabling continuous improvement and adaptation to new agronomic information.

The retrieval pipeline also integrates a reranking step using the Cohere Re-ranker, as documented in LangChain's retriever integrations. This reranker further refines the initial set of retrieved documents by considering relevance scores, leading to improved answer quality.

4.5.4.3 Response Composition

The final system adopts a chained approach composed of the following stages:

1. **Query Transformation:** The original user query is transformed using an LLM to optimize its effectiveness for retrieval. This step helps reformulate potentially

vague or incomplete user questions into more precise search queries.

2. **Document Retrieval:** The transformed query is then used to retrieve the most relevant documents from the FAISS index, as described above.

3. **Response Generation:** Retrieved documents are formatted and incorporated into a prompt passed to an LLM for final response generation. The specialized RAG prompt template:

    - Positions the LLM as an expert in plant disease diagnosis.
    - Explicitly instructs the model to base its answers on the retrieved documents and to acknowledge when information is insufficient.
    - Encourages the inclusion of source citations derived from the retrieved materials.
    - Passes both the original user query and the retrieved documents in a structured format to maximize context awareness.

The entire workflow is orchestrated using **LangGraph**, which manages the state transitions between the query transformation, retrieval, and response generation steps. This graph-based orchestration enables modularity and clear separation of responsibilities within the RAG pipeline.

Overall, this implementation creates a flexible and robust RAG system capable of enhancing retrieval effectiveness through query transformation and generating contextually grounded responses that reduce hallucination and improve user trust.

4.5.5 Synthetic Dataset Generation for Evaluation

To robustly evaluate our RAG implementations, we constructed a synthetic test dataset grounded in the system's agronomic knowledge base. This approach allows the evaluation to be automated and repeatable without relying on extensive manual annotations.

We adopted the testset generation strategy proposed by the Retrieval-Augmented Generation Assessment Suite (RAGAS), which consists of the following steps:

1. **Document Loading:** The reference agronomic documents (books and expert materials) were ingested using the `DirectoryLoader` class from the `langchain_community` module.

2. **LLM Selection:** A large language model (e.g., GPT-4) was chosen to generate question–answer pairs. The LLM was wrapped using the `LangchainLLMWrapper` to integrate smoothly with the generation pipeline.

3. **Prompting for Q&A Generation:** The LLM was prompted to create diverse and representative question–answer pairs based on the loaded documents. The generated questions ranged from simple factual checks to more complex, multi-hop, or edge-case scenarios.

4. **Dataset Structuring:** The generated dataset included columns such as `question`, `ground_truth` (expected answer), and `contexts` (reference sections supporting the answer). This dataset formed the basis for systematic evaluation.

This synthetic approach ensured that the evaluation set was specifically tailored to the actual knowledge encoded in the system, while maintaining flexibility and reproducibility. While synthetic datasets enable controlled and scalable evaluation, they may not fully capture the linguistic variability and noise present in real farmer interactions, representing a potential limitation.

4.5.6 RAGAS Evaluation Framework

The three RAG approaches (semantic-only, hybrid, and query-transformed) were evaluated using the RAGAS framework. RAGAS provides a comprehensive, reference-free evaluation of RAG pipelines, focusing on both retrieval and generation quality without requiring human-labeled ground truth annotations.

4.5.6.1 Evaluation Metrics

We used the `evaluate()` function from RAGAS to assess each RAG approach based on a set of core metrics:

- **Context Precision**: Measures the proportion of retrieved documents that are actually relevant to the question.
- **Context Recall**: Evaluates whether all essential information needed to answer the question is present in the retrieved documents.
- **Context Entities Recall**: Assesses whether key entities mentioned in the question and answer are present in the retrieved contexts.
- **Noise Sensitivity**: Examines how susceptible the retrieval process is to irrelevant or misleading information.
- **Response Relevancy**: Checks if the generated response appropriately addresses the question.
- **Faithfulness**: Evaluates whether the response remains grounded in the retrieved contexts and does not hallucinate information.

Each RAG variant was tested against the same set of 100 synthetic questions and expected answers, containing the same literature database. The metrics provided a granular analysis of retrieval quality, answer accuracy, and factual alignment, supporting a fair comparative study.

These comprehensive evaluations guided the selection and refinement of the final RAG approach, helping to achieve higher reliability and trustworthiness in the plant diagnosis system.

4.5.7 Model Context Protocols

4.5.7.1 Definition and Origin

The *Model Context Protocol* (MCP) is a crucial open standard communication protocol that facilitates structured interaction between large language models (LLMs) and external systems, tools, and data sources. Introduced by Anthropic, MCP is publicly available as an open standard (ANTHROPIC, -). It is often described as a “universal connector,” likened to a “USB-C port for AI applications,” because it standardizes how different systems expose resources (data, functions, workflows) to LLMs, eliminating the need for bespoke code for each integration(MCP, -).

4.5.7.2 Core Architecture and Mechanism

MCP employs a client-server architecture, where:

- An MCP “Host” application (e.g., an LLM-powered agent or front-end) contains one or more MCP clients. Each client connects to an MCP Server, which provides a set of *tools*, *resources*, and *prompts* for the LLM to invoke or consume (RAY, 2025).
- The protocol defines the exchange of messages between client and server, using formats like JSON-RPC, to manage capability discovery, context provision, tool invocation, and result return (MCP, -).

In practical terms, the LLM (or its host application) queries the MCP server for available actions and data, selects the appropriate tool/resource, and receives the output through a standardized interface. This allows the model to extend beyond pure text generation, effectively gaining the ability to perform actions like running APIs, accessing databases, or calling functions via MCP.

4.5.7.3 MCP as an Extension of LLM Capabilities

MCP enables LLMs to extend their capabilities beyond text generation by dynamically interacting with external systems. This interaction allows LLMs to perform tasks such as data retrieval, computation, and workflow orchestration in real-time. For example:

- When an LLM identifies the need for data from a relational database, it uses MCP to invoke a “query DB” tool provided by an MCP server.
- If computation or an API call is required, the LLM triggers the tool, awaits the result, and integrates the result into its output.
- For consulting a knowledge base or orchestrating workflows, MCP enables seamless workflow orchestration through tools and chained calls.

MCP abstracts the complexities of each tool or data source (file system, HTTP API, database, etc.), providing a standardized interface for integration without bespoke logic.

This architecture simplifies the development of agent-style AI systems and enhances the LLM's ability to interact with diverse environments(RAY, 2025).

4.5.7.4 Relevance for Our System

In our plant-disease diagnosis architecture, adopting MCP (or a similar modular protocol) enhances our system's capability to integrate diverse data sources and tools, improving diagnostic accuracy and efficiency. Based on the modular architecture, several specialized MCPs were designed to handle different aspects of plant disease diagnosis:

- **Text Diagnosis MCP**: Uses LLMs and the knowledge base to analyze textual symptom descriptions and provide possible disease matches with confidence levels.
- **Image Diagnosis MCP**: Employs deep learning models (specifically ResNet-50 and EfficientNetB3) fine-tuned on plant disease image datasets to analyze uploaded photos.
- **Treatment Recommendation MCP**: Retrieves and formats appropriate treatment protocols based on the identified diseases and severity levels.

4.5.8 Database Layer

The system implements a two-tier database architecture:

- **Vector storage**: Embeddings are stored in PostgreSQL with the `pgvector` extension, enabling efficient similarity searches across the agricultural knowledge base.
- **Session management**: Conversation history and user context are maintained in a separate database layer that implements both short-term session memory (active conversations) and long-term memory (previous diagnoses and user preferences).

To address the challenge of outdated context, the system implements a time-decay mechanism that progressively reduces the relevance weight of older context entries. Additionally, the database periodically prunes stale entries that exceed a configurable retention threshold, maintaining optimal performance while preserving important historical data.

4.5.9 Cloud and Containerization

Each component is containerized and orchestrated with Docker Compose for development and testing environments. Environment variables manage configuration settings, while Docker volumes ensure data persistence across container restarts. Health checks are implemented to provide resilience and automatic recovery from failures.

For production deployment, the architecture is designed to be compatible with Kubernetes, allowing for future migration as demand scales. The modular design facilitates horizontal scaling of individual components, particularly the computationally intensive image analysis and RAG retrieval services, without requiring a complete system redesign.

4.6 DETAILED DESCRIPTION OF RAG COMPONENT

4.6.1 Motivation and Design Choices

The RAG component was designed to address the limitations of purely generative LLM outputs, particularly the risk of hallucinations. By grounding responses in a curated agronomic knowledge base, RAG ensures factual accuracy and enhances user trust.

4.6.2 Architecture and Workflow

The RAG component follows a pipeline architecture comprising several key stages:

1. **Query transformation**: The user's input is analyzed and transformed into an optimized search query, expanding agricultural terminology and identifying key disease indicators.

2. **Retrieval**: The system performs parallel retrieval operations:
    - Semantic search using `FAISS` vector database with OpenAI embeddings
    - Keyword-based search for specific disease names, crops, or technical terms
    - Metadata filtering by crop type, region, and severity when this context is available
3. **Reranking**: Retrieved documents are scored and filtered based on relevance, recency, and source reliability. This step particularly improved precision in the final Agentic RAG prototype.
4. **Context assembly**: The highest-ranked documents are combined with user query and conversation history to form a comprehensive context window.
5. **Response generation**: The LLM generates a response based on the assembled context, with explicit instructions to cite sources and acknowledge uncertainty when information is incomplete.

This workflow evolved substantially across the three prototypes, with the final implementation incorporating feedback loops for query refinement and explicit handling of edge cases where insufficient information is available in the knowledge base.

The RAG component operates in the following stages:

1. **Query Encoding**: User queries are encoded into embeddings using an OpenAI embedding model, ensuring semantic representation.
2. **Similarity Search**: The encoded query is compared against stored document embeddings using cosine similarity within a pgvector-enabled PostgreSQL database.
3. **Chunk Retrieval**: The system retrieves the top-k most similar text chunks (e.g., $k = 5$), each representing approximately 1000 characters with 200-character overlaps. These chunks contain domain knowledge sourced from agronomy books and technical manuals.
4. **Contextual Augmentation**: The retrieved chunks are concatenated and incorporated into the LLM prompt as additional context.
5. **Answer Generation**: The LLM generates a final, grounded response using both the retrieved information and its inherent language capabilities.

4.6.3 Embedding and Storage Details

Chunks were generated from OCR-processed agronomic literature. After cleaning, text was divided into overlapping segments to maintain context continuity. Embeddings were generated using OpenAI's embedding API and stored in a PostgreSQL database with pgvector extension, chosen for its efficient vector indexing and integration simplicity.

4.6.4 Future Improvements

Potential improvements include hybrid retrieval strategies combining semantic and keyword-based search, dynamic chunk sizing based on query complexity, and re-ranking methods to further optimize relevance.

4.7 DETAILED DESCRIPTION OF IMAGE DIAGNOSIS MCP

4.7.1 Model Architecture and Dataset

The Image Diagnosis MCP uses a convolutional neural network (CNN) architecture, such as ResNet or EfficientNet, chosen for their balance between accuracy and computational efficiency. The model was fine-tuned on a dataset comprising labeled images of plant leaves exhibiting various disease symptoms.

To improve generalization, extensive data augmentation techniques were applied, including random rotations, horizontal and vertical flips, brightness adjustments, and zoom variations.

4.7.2 Integration with Text-Based Diagnosis

The Image MCP outputs predicted disease classes with associated confidence scores. These results are then cross-referenced with text-based diagnoses from the RAG component. In cases where both modalities agree, confidence in the final recommendation increases. In conflicting cases, the system may either request additional user input or present a combined explanation outlining uncertainties.

4.7.3 Limitations and Potential Biases

Despite the robustness of the proposed methodology, several limitations exist. The reliance on synthetic evaluation data, while practical, may not fully capture real-world query variability. Additionally, the system's dependence on LLMs for query transformation and question generation introduces potential model biases. Finally, using FAISS as a local vector store may constrain scalability in distributed or large-scale deployments. Future work may explore incorporating real user query logs and evaluating on larger, multi-region deployments to address these constraints.

4.7.4 Future Work

Future enhancements may involve integrating multi-label classification to handle co-occurring diseases, expanding the dataset with new crops and regions, and exploring explainable AI techniques (e.g., Grad-CAM) to visually highlight affected areas on images.

4.8 MEMORY AND CONTEXTUAL REASONING

Session memory is handled through vector embeddings in pgvector, enabling context retention over multiple user interactions. This approach improves coherence, avoids repetitive queries, and supports user-centric experiences.

4.9 SECURITY AND PRIVACY CONSIDERATIONS

API keys and sensitive data are managed via environment variables and not hard-coded. User data is stored temporarily and not shared externally, ensuring compliance with privacy best practices. Deployment options include secure, isolated cloud environments.

4.10 LESSONS LEARNED AND CHALLENGES

The evolution from simple RAG-only workflows to multi-agent architectures and finally to a streamlined single-agent design revealed trade-offs between flexibility, robustness, and maintainability.

While multi-agent coordination improved certain capabilities, it introduced maintenance complexity. The final unified agent design, grounded in RAG, improved robustness while reducing operational overhead.

4.11 CHAPTER CONCLUSION

This chapter presented the iterative development of the LIMMO system, detailing the transition from a simple RAG-based workflow to a multi-agent architecture and finally to a robust, modular single-agent design. Each evolution reflected practical lessons learned through testing and qualitative feedback rather than strict quantitative metrics. The final architecture effectively balances technical sophistication with operational maintainability, providing a powerful, accurate, and explainable diagnostic assistant. This foundation supports the system's intended deployment in field conditions, directly serving agronomists and farmers through accessible and transparent AI-driven recommendations.

5 RESULTS AND DISCUSSION

5.1 FRAMEWORK SELECTION RESULTS

The comparative analysis of frameworks (CrewAI, Phidata, and LangChain/LangGraph) yielded valuable insights that guided architectural decisions. While CrewAI offered convenient role-based agent design, its higher-level abstractions limited fine-grained control over the RAG pipeline components. Similarly, Phidata provided excellent out-of-the-box RAG capabilities but became constraining when implementing specialized diagnostic workflows.

LangChain with LangGraph emerged as the optimal choice, offering several advantages:

- Greater modularity for customizing retrieval and reranking logic
- Extensive ecosystem of connectors for vector stores and embedding models
- Visual debugging capabilities for complex workflows
- Mature community support and documentation

This selection proved critical in facilitating the rapid prototyping approach while maintaining the flexibility needed for specialized agricultural diagnostics.

5.2 PROTOTYPE ITERATIONS AND IMPROVEMENTS

Three distinct prototypes were developed and evaluated, each representing a significant architectural evolution:

5.2.1 First Prototype: Semantic-Only RAG

The initial implementation relied solely on basic semantic retrieval using OpenAI embeddings and FAISS. This approach, while straightforward to implement, showed significant limitations:

- High hallucination rate when knowledge gaps existed.

- Limited precision in retrieving contextually relevant information.
- Difficulty handling ambiguous agricultural terminology.
- Over-reliance on embedding quality and database coverage.

RAGAS evaluation metrics confirmed these observations, with the Semantic-Only RAG approach scoring lowest across all measured dimensions (Context Precision: 0.72, Faithfulness: 0.63).

5.2.2 Second Prototype: Hybrid RAG

The second iteration introduced a hybrid RAG mechanism combining semantic similarity with keyword-based retrieval. This adjustment significantly improved factual consistency and reduced hallucinations. The integration of `LangChain` and `LangGraph` allowed for modular experimentation and enhanced control over agent workflows.

Measurable improvements included:

- Increased Context Precision from 0.72 to 0.81.
- Improved Context Recall from 0.65 to 0.78.
- Reduced Noise Sensitivity from 0.30 to 0.22.
- Enhanced Faithfulness from 0.63 to 0.77.

5.2.3 Final Prototype: Agentic RAG with MCPs

In the final prototype, the system was further optimized by introducing specialized Model Context Protocols (MCPs) for text diagnosis, image diagnosis, and treatment recommendation, along with implementing query transformation and response composition techniques. The Agentic RAG approach with MCPs showed notable improvements across all evaluation metrics:

- Context Precision increased to 0.89.
- Context Recall reached 0.85.

- Entity Recall improved to 0.82.
- Noise Sensitivity decreased to 0.15.
- Response Relevancy rose to 0.88.
- Faithfulness achieved 0.86.

These improvements translated to tangible user experience benefits, with faster response times and more accurate diagnoses.

5.3 QUANTITATIVE RESULTS

5.3.1 RAG Evaluation Results

The RAGAS evaluation framework provided comprehensive insights into the performance of each RAG variant. Table 3 shows the comparative metrics across all three approaches:

Table 3 – Comparative RAG evaluation results on 100 synthetic questions using RAGAS metrics

Approach	Context Prec.	Context Recall	Entity Recall	Noise Sens.	Relevancy	Faithfulness
Semantic-Only RAG	0.72	0.65	0.58	0.30	0.68	0.63
Hybrid RAG	0.81	0.78	0.73	0.22	0.79	0.77
Agentic RAG	0.89	0.85	0.82	0.15	0.88	0.86

These metrics demonstrate a clear progression in retrieval and generation quality across prototypes, with the final agentic approach showing substantial improvements in all dimensions.

5.3.2 Accuracy of Diagnosis

The system was further evaluated on a set of 50 real plant disease cases, including both text-based symptom descriptions and image submissions. The overall diagnostic accuracy showed consistent improvement across iterations:

- **First prototype (Semantic-Only RAG)**: Approximately 60% correct diagnostic suggestions.

- **Second prototype (Hybrid RAG)**: Approximately 75% correct diagnostic suggestions.
- **Final prototype (Agentic RAG with MCPs)**: Over 90% correct diagnostic suggestions.

These improvements highlight the value of integrating domain-specific retrieval strategies, query transformation, and multimodal input processing to achieve higher diagnostic reliability.

5.3.3 Response Time and Resource Utilization

System response times were measured as a critical user experience metric. Despite the increased sophistication of the agentic RAG approach, performance optimizations yielded significant improvements:

- **Text-based diagnoses**: Decreased from >15 seconds to approximately 5 seconds.
- **Image-based diagnoses**: Reduced from >25 seconds to approximately 10 seconds.
- **Memory utilization**: Decreased by approximately 30% in the final prototype.

These efficiency gains were achieved through optimized embedding generation, strategic caching, and streamlining the agent workflow from multi-agent to single-agent with modular components.

5.4 QUALITATIVE FEEDBACK

Preliminary feedback was collected from a panel of five expert agronomists and ten regular users (smallholder farmers). The feedback highlighted several key strengths of the final system:

- **Contextual relevance**: Experts noted that responses demonstrated appropriate regional and crop-specific knowledge.

- **Multimodal flexibility**: Users appreciated the ability to seamlessly switch between text descriptions and image inputs.
- **Factual grounding**: The system's tendency to cite specific sources was noted as enhancing trust and credibility.
- **Appropriate uncertainty**: When facing ambiguous inputs, the system appropriately communicated uncertainty rather than making overconfident diagnoses.

Particularly noteworthy was the feedback regarding hallucination reduction. Expert reviewers identified only 3 instances of minor factual inconsistencies across 50 test cases, compared to 17 such instances in the first prototype.

5.5 DISCUSSION

5.5.1 RAG Approach Effectiveness

The results demonstrate the clear superiority of the Agentic RAG approach with specialized MCPs over both the Semantic-Only RAG and Hybrid RAG strategies. The progression from basic vector similarity to sophisticated query transformation and contextual augmentation yielded measurable improvements in retrieval precision, factual consistency, and diagnostic accuracy. Notably, the final Agentic RAG system was further enhanced with external data access capabilities through the Tavily web search API and Embrapa API integration, allowing it to provide accurate information even when the local knowledge base was insufficient—a significant advantage over the earlier prototypes which relied solely on locally stored knowledge.

The synthetic evaluation dataset approach, while potentially limiting in its representation of real-world query diversity, provided valuable benchmarks for systematic comparison. These findings align with recent research showing that LLM-driven query transformation can significantly enhance retrieval performance in domain-specific applications.

5.5.2 Architectural Trade-offs

The evolution from multi-agent to single-agent architecture with MCPs represented a critical design decision. While the multi-agent approach offered theoretical benefits in specialization, the practical challenges in agent coordination, state management, and orchestration complexity outweighed these advantages. The single-agent architecture with modular MCPs delivered superior performance with reduced complexity, supporting the principle that simpler architectures often yield more robust and maintainable systems.

This finding contradicts some current trends toward complex multi-agent systems and suggests that in specialized domains like agricultural diagnostics, architectural simplicity with focused modularity may be preferable.

5.5.3 Limitations

Despite the promising results, several limitations should be acknowledged:

- The reliance on synthetic evaluation data (100 questions) may not fully capture the diversity and complexity of real-world queries that farmers and agronomists would generate.
- The system's dependence on LLMs introduces potential model biases that could impact agricultural contexts differently across regions.
- Scalability testing under high-concurrency scenarios was limited and requires further investigation.
- The local `FAISS` vector store implementation may face limitations in distributed deployments.

Additionally, the image diagnosis MCP demonstrated a robust architecture with built-in redundancy. It primarily utilizes the CultivAI API for initial image analysis, but seamlessly falls back to GPT-based image interpretation when the primary extraction fails or when users question the accuracy of results. This dual-model approach significantly improved resilience and accuracy across diverse plant varieties and image

qualities, though its effectiveness still remains partially constrained by the quality and diversity of the training datasets, which may not represent all regional crop varieties and disease manifestations.

5.6 KEY INSIGHTS

- **RAG Strategy Impact**: The transition from purely Semantic-Only RAG to an Agentic RAG approach with query transformation yielded the most significant improvements in diagnosis accuracy and factual consistency.
- **Architectural Simplification**: The move from complex multi-agent systems to a streamlined single-agent with modular MCPs improved both performance and maintainability.
- **Framework Selection Importance**: The choice of `LangChain` with `LangGraph` proved critical in enabling rapid experimentation while maintaining sufficient control over RAG pipeline components.
- **Multimodal Synergy**: The integration of text and image processing created a system greater than the sum of its parts, mimicking the natural diagnostic workflow of human experts.
- **Evaluation Methodology**: The RAGAS framework provided valuable, multi-dimensional insights into RAG performance that would not have been captured by simple accuracy metrics alone.

5.7 FUTURE DIRECTIONS

Building on these findings, several promising research and development directions emerge:

- Expanding data sources and knowledge resources to achieve better coverage of plant care issues beyond disease diagnosis, including nutrient deficiencies, pest management, and cultivation best practices

- Conducting large-scale real-world testing with farmers and agronomists to improve system responsiveness and accuracy in practical field conditions
- Developing a disease/issues tracking system that derives insights from platform usage and enables geographic and temporal tracking of plant disease outbreaks, creating an early warning system for agricultural stakeholders
- Refining the self-correction mechanisms, as the current implementation requires users to recognize and challenge incorrect diagnoses—a capability that varies widely among users with different expertise levels
- Investigating dynamic chunk sizing and context window optimization to further improve retrieval efficiency
- Implementing more sophisticated reranking strategies to enhance response quality, particularly for specialized agricultural terminology
- Exploring distributed vector store architectures for improved scalability in production environments
- Incorporating continuous learning mechanisms to allow the system to improve from user interactions and feedback
- Evaluating the system's performance across different cultural and linguistic contexts in agricultural settings
- Differentiating capabilities from general-purpose tools (like ChatGPT, Gemini, or Perplexity AI) by focusing on specialized agricultural knowledge that leverages domain-specific data rather than relying solely on web search for broader topics

These directions would address current limitations while building on the strong foundation established by the iterative prototyping approach.

6 CONCLUSION

6.1 SUMMARY OF CONTRIBUTIONS

This dissertation has presented a comprehensive investigation into the design, implementation, and evaluation of a multimodal conversational system for plant disease diagnosis. The work makes several significant contributions to the field of agricultural artificial intelligence:

- **A systematic evaluation of RAG approaches for agricultural diagnostics**, progressing from simple semantic retrieval to hybrid approaches and finally to a sophisticated agentic RAG system with modular components. The quantitative RAGAS evaluation demonstrated substantial improvements across all metrics, with faithfulness increasing from 0.63 to 0.86 and context precision from 0.72 to 0.89.
  - **Systematic RAG evaluation**: The research demonstrated a methodical progression from basic Semantic-Only RAG to Hybrid RAG and finally Agentic RAG approaches, providing quantifiable evidence of performance improvements across multiple dimensions.
  - **Framework analysis and selection**: Through comparative evaluation of emerging generative AI frameworks (`CrewAI`, `Phidata`, and `LangChain/Graph`), the research established criteria for selecting optimal architectures for agricultural applications.
  - **Architectural simplification**: The successful transition from complex multi-agent systems to a streamlined single-agent with modular MCPs offers a blueprint for maintainable AI systems in resource-constrained domains.
- **A multimodal agricultural diagnostic system** that integrates text, image, and audio inputs through a unified framework, closely mimicking the natural diagnostic workflow of human agronomists while maintaining system coherence and factual accuracy.

- **A robust evaluation methodology** combining synthetic question generation, RAGAS metrics, and expert validation to provide multi-dimensional assessment of agricultural AI systems beyond simple accuracy measurements.

The iterative development process demonstrated that careful framework selection and architectural decisions significantly impact not only technical performance but also development velocity and system maintainability. The final prototype achieved over 90% diagnostic accuracy while reducing response times by more than 50% compared to initial implementations.

6.2 KEY FINDINGS

6.2.1 RAG Implementation Insights

The comprehensive evaluation of different RAG approaches yielded several important insights for domain-specific knowledge systems:

- Query transformation using domain-aware LLMs significantly improves retrieval precision in agricultural contexts, where terminology may be ambiguous or regionally varied.
- Hybrid retrieval combining semantic similarity with keyword matching outperforms pure vector-based approaches, particularly for technical agricultural terms and specific disease names.
- The agentic RAG approach with explicit state management demonstrates superior faithfulness (0.86 vs 0.63) compared to simpler implementations, suggesting that sophisticated retrieval orchestration is crucial for reducing hallucinations in complex domains.
- The RAGAS evaluation framework reveals nuanced performance differences that would be missed by simple accuracy metrics, highlighting the importance of multi-dimensional evaluation for RAG systems.

6.2.2 Architectural Findings

The evolution from multi-agent to single-agent architecture yielded several counter-intuitive findings:

- **RAG implementation**: The progressive refinement of retrieval mechanisms—from basic Semantic-Only RAG to sophisticated query transformation with Hybrid RAG approaches yielded substantial improvements in diagnosis accuracy. The final Agentic RAG implementation with adaptive query planning reduced hallucinations by approximately 80% compared to the initial prototype.
- **Architectural findings**: The research challenged the common assumption that multi-agent architectures are inherently superior for complex tasks. By implementing a streamlined single-agent system with specialized MCPs for text diagnosis, image diagnosis, and treatment recommendation, the project achieved both performance gains and improved maintainability. `LangGraph`'s graph-based workflow control provided the right balance of flexibility and structure.

User feedback strongly validated these architectural choices, with both expert agronomists and smallholder farmers noting improved response relevance, reduced hallucinations, and more intuitive multimodal interactions in the final system.

6.3 LIMITATIONS

Despite the significant advances demonstrated, several important limitations must be acknowledged:

- **Evaluation constraints**: While the RAGAS framework provided valuable multi-dimensional metrics, the evaluation relied heavily on synthetic test cases (100 questions). Real-world performance may vary with the 50 real test cases, particularly with unusual or ambiguous disease presentations.
- **Model dependencies**: The system's performance is tied to the underlying LLMs and vision models, introducing external dependencies that may affect long-term reliability and consistency.

- **Infrastructure requirements**: Despite optimizations, the system still requires consistent internet connectivity and moderate computing resources, potentially limiting deployment in extremely remote agricultural settings.
- **Regional adaptability**: While effort was made to include diverse agricultural contexts, the knowledge base has stronger coverage of major commercial crops than region-specific or lesser-known varieties.

Additionally, the current implementation does not incorporate continuous learning capabilities, meaning that knowledge updates require manual intervention rather than occurring organically through system usage.

6.4 FUTURE WORK

Building on the foundation established in this research, several promising directions for future work emerge:

- **Enhanced retrieval strategies**: Future work should explore more sophisticated retrieval methods, including hierarchical indexing and domain-specific embedding models trained specifically on agricultural terminology.
- **Offline capabilities**: Developing lightweight, offline-capable models would address connectivity limitations in remote areas.
- **Longitudinal studies**: Extended field testing across multiple growing seasons would provide more comprehensive validation of system effectiveness and reliability.
- **Knowledge integration**: Establishing automated pipelines for incorporating new research findings and disease reports would ensure the system remains current with emerging agricultural challenges.
- **Distributed vector storage**: Exploring more scalable approaches to vector database management would support the system's growth beyond the current PostgreSQL with `pgvector` implementation.

These directions would address current limitations while expanding the system's practical utility across diverse agricultural contexts.

6.5 FINAL REMARKS

This research demonstrates that well-designed agricultural AI systems combining retrieval-augmented generation, multimodal capabilities, and careful architectural design can achieve high levels of diagnostic accuracy while maintaining factual consistency. The systematic evaluation of different RAG strategies and architectural approaches provides valuable insights for developing domain-specific knowledge systems beyond agriculture.

By prioritizing framework flexibility, architectural simplicity, and robust evaluation methodologies, this work offers a blueprint for creating AI systems that genuinely empower agricultural communities rather than introducing new dependencies or complexities. The evolution from complex multi-agent systems to streamlined, modular architectures highlights that sophistication in AI does not necessarily require complexity, thoughtful simplification often yields superior results.

Ultimately, this research contributes to the broader goal of making agricultural expertise more accessible, helping smallholder farmers and agricultural professionals make better-informed decisions through AI systems that are factual, contextually aware, and practically useful. As these technologies continue to evolve, their potential to democratize agricultural knowledge and support sustainable farming practices worldwide will only increase.

BIBLIOGRAPHY

ACP. *Welcome - Agent Communication Protocol*. 2025. [Online; accessed 2025-06-12]. Available at: <https://agentcommunicationprotocol.dev/introduction/welcome>.

ALI, E. *Rapid and Accurate Disease Diagnosis as a Key Component to Successful Plant Disease Management | Extension*. 2022. [Online; accessed 2025-10-24]. Available at: <https://extension.unh.edu/blog/2022/02/rapid-accurate-disease-diagnosis-key-component-successful-plant-disease-management>.

ANTHROPIC. *Introducing the Model Context Protocol*. –. [Online; accessed 2025-10-23]. Available at: <https://www.anthropic.com/news/model-context-protocol>.

ANTHROPIC. *Introducing the Model Context Protocol \ Anthropic*. 2024. [Online; accessed 2025-06-12]. Available at: <https://www.anthropic.com/news/model-context-protocol>.

ARTETXE, M.; DU, J.; GOYAL, N.; ZETTLEMOYER, L.; STOYANOV, V. On the role of bidirectionality in language model pre-training. *arXiv preprint arXiv:2205.11726*, 2022.

AWASTHI, Y. Press “a”’ for artificial intelligence in agriculture: A review. *JOIV: International Journal on Informatics Visualization*, v. 4, n. 3, p. 112–116, 2020.

BAIJU, B.; KIRUPANITHI, N.; SRINIVASAN, S.; KAPOOR, A.; MATHIVANAN, S. K.; SHAH, M. A. Robust crw crops leaf disease detection and classification in agriculture using hybrid deep learning models. *Plant Methods*, Springer, v. 21, n. 1, p. 18, 2025.

BELOEV, I.; KINANEVA, D.; GEORGIEV, G.; HRISTOV, G.; ZAHARIEV, P. Artificial intelligence-driven autonomous robot for precision agriculture. *Acta Technologica Agriculturae*, v. 24, n. 1, p. 48–54, 2021.

BERA, A.; BHATTACHARJEE, D.; KREJCAR, O. An attention-based deep network for plant disease classification. *Machine Graphics & Vision*, v. 33, n. 1, p. 47–67, 2024.

BERA, A.; BHATTACHARJEE, D.; KREJCAR, O. Pnd-net: plant nutrition deficiency and disease classification using gnn. *Scientific Reports*, v. 14, p. 15537, 2024.

BREZULEANU, M.-M.; UNGUREANU, E.; ZAHARIA, R. S.; BREZULEANU, C.-O. Studies on the sustainable impact of artificial intelligence in technical and agricultural university education: Benefits challenge and future directions. *Scientific Papers Series Management, Economic Engineering in Agriculture & Rural Development*, v. 25, n. 1, 2025.

CHEN, L.; CHEN, Z.; ZHANG, Y.; LIU, Y.; OSMAN, A. I.; FARGHALI, M.; HUA, J.; AL-FATESH, A.; IHARA, I.; ROONEY, D. W. et al. Artificial intelligence-based solutions for climate change: a review. *Environmental Chemistry Letters*, Springer, v. 21, n. 5, p. 2525–2557, 2023.

CHEN, W.; LIU, X.; WANG, J. Graph neural networks for plant disease classification: A novel approach. *IEEE Transactions on Artificial Intelligence*, 2023.

CHEN, Z.; WANG, G.; LV, T.; ZHANG, X. Using a hybrid convolutional neural network with a transformer model for tomato leaf disease detection. *Agronomy*, v. 14, n. 4, 2024. ISSN 2073-4395. Available at: <https://www.mdpi.com/2073-4395/14/4/673>.

CHEN, Z.; WANG, G.; LV, T.; ZHANG, X. Using a hybrid convolutional neural network with a transformer model for tomato leaf disease detection. *Agronomy*, MDPI, v. 14, n. 4, p. 673, 2024.

CISCO. *Outshift | From concept to code: AGNTCY'S Internet of Agents is now on GitHub*. 2025. [Online; accessed 2025-06-12]. Available at: <https://outshift.cisco.com/blog/agntcy-internet-of-agents-is-on-github>.

CREWAI. –. [Online; accessed 2025-10-23]. Available at: <https://crewai.com/>.

DEVLIN, J.; CHANG, M.-W.; LEE, K.; TOUTANOVA, K. N. Bert: Pre-training of deep bidirectional transformers for language understanding. In: . [s.n.], 2018. Available at: <https://arxiv.org/abs/1810.04805>.

DEVLIN, J.; CHANG, M.-W.; LEE, K.; TOUTANOVA, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *NAACL*, 2019.

DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T.; DEHGHANI, M.; MINDERER, M.; HEIGOLD, G.; GELLY, S.; USZKOREIT, J.; HOULSBY, N. An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations*, 2021.

FERENTINOS, K. P. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, v. 145, p. 311–318, 2018.

FRANTAR, E. et al. Gptq: Accurate post-training quantization for generative pretrained transformers. *ICLR*, 2023.

GHOSH, H.; RAHAT, I. S.; EMON, M. M. R.; MASHRAFI, M. J.; TANZIN, M. A. A. A.; MOHANTY, S. N.; KANT, S. Advanced neural network architectures for tomato leaf disease diagnosis in precision agriculture. *Discover Sustainability*, Springer, v. 6, n. 1, p. 312, 2025.

GOH, E.; GALLO, R.; HOM, J.; STRONG, E.; WENG, Y.; KERMAN, H.; COOL, J. A.; KANJEE, Z.; PARSONS, A. S.; AHUJA, N.; HORVITZ, E.; YANG, D.; MILSTEIN, A.; OLSON, A. P. J.; RODMAN, A.; CHEN, J. H. Large language model influence on diagnostic reasoning: A randomized clinical trial. *JAMA Network Open*, v. 7, n. 10, p. e2440969–e2440969, 10 2024. ISSN 2574-3805. Available at: <https://doi.org/10.1001/jamanetworkopen.2024.40969>.

GOMEZ-GUTIERREZ, S. V.; GOODWIN, S. B. Loop-mediated isothermal amplification for detection of plant pathogens in wheat (triticum aestivum). *Frontiers in plant science*, Frontiers Media SA, v. 13, p. 857673, 2022.

GONZALES, K. K. C.; DIOSES, I. A. M. Efficientnet convolutional neural network approach in classifying multiple tomato diseases. In: IEEE. *2024 IEEE 12th Conference on Systems, Process & Control (ICSPC)*. [S.l.], 2024. p. 257–262.

GONZÁLEZ-BRIONES, A.; FLOREZ, S. L.; CHAMOSO, P.; CASTILLO-OSSA, L. F.; CORCHADO, E. S. Enhancing plant disease detection: Incorporating advanced cnn architectures for better accuracy and interpretability. *International Journal of Computational Intelligence Systems*, Springer, v. 18, n. 1, p. 120, 2025.

GOOGLE. *Announcing the Agent2Agent Protocol (A2A) - Google Developers Blog*. 2025. [Online; accessed 2025-06-12]. Available at: <https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/>.

GOURABATHINA, A.; GERYCH, W.; PAN, E.; GHASSEMI, M. The medium is the message: How non-clinical information shapes clinical decisions in llms. In: *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery, 2025. (FAccT '25), p. 1805–1828. ISBN 9798400714825. Available at: <https://doi.org/10.1145/3715275.3732121>.

GUO, X.; FENG, Q.; GUO, F. Cmtnet: a hybrid cnn-transformer network for uav-based hyperspectral crop classification in precision agriculture. *Scientific Reports*, Nature Publishing Group UK London, v. 15, n. 1, p. 12383, 2025.

GUU, K.; LEE, K.; TUNG, Z.; PASUPAT, P.; CHANG, M.-W. Realm: Retrieval-augmented language model pre-training. In: *International Conference on Machine Learning*. [S.l.: s.n.], 2020. p. 3929–3938.

HENDRYCKS, D.; BURNS, C.; BASART, S.; ZOU, A.; MAZEIKA, M.; SONG, D.; STEINHARDT, J. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*, 2020.

HENDRYCKS, D. et al. Measuring massive multitask language understanding. *ICLR*, 2021.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural Computation*, MIT press, v. 9, n. 8, p. 1735–1780, 1997.

HU, B.; JIANG, W.; ZENG, J.; CHENG, C.; HE, L. Fotca: hybrid transformer-cnn architecture using afno for accurate plant leaf disease image recognition. *Frontiers in Plant Science*, Frontiers Media SA, v. 14, p. 1231903, 2023.

IBGE. *Censo Agrícola 2017*. 2017. <https://censoagro2017.ibge.gov.br/templates/censo/_agro/resultadosagro/index.html>. [Acessado em 1 de Março de 2021].

IZACARD, G.; GRAVE, E. Leveraging passage retrieval with generative models for open domain question answering. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. [S.l.: s.n.], 2020. p. 8749–8769.

JAHIN, M. A.; SHAHRIAR, S.; MRIDHA, M. F.; HOSSEN, M. J.; DEY, N. Soybean disease detection via interpretable hybrid cnn-gnn: Integrating mobilenetv2 and graphsage with cross-modal attention. *arXiv preprint arXiv:2503.01284*, 2025.

JI, Z.; LEE, N.; FRIESKE, R.; YU, Y.; AL. et. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 2023.

JINDAL, M.; KAUR, K. Enhancing agricultural sustainability through ai-powered image processing: Review study on plant disease detection. *International Journal of Scientific Research in Science and Technology*, 2024. Review of AI systems integrating imaging and LLMs in agriculture.

JOHNSON, J. et al. Faiss: Facebook ai similarity search. *arXiv preprint arXiv:1702.08734*, 2019.

KANAPIYA, A.; AMANBAYEVA, U.; TULEGENOVA, Z.; ABASH, A.; ZHANGAZIN, S.; DYUSSEMBAYEV, K.; MUKIYANOVA, G. Recent advances and challenges in plant viral diagnostics. *Frontiers in Plant Science*, Frontiers Media SA, v. 15, p. 1451790, 2024.

KAZEMI, M.; FATEMI, B.; BANSAL, H.; PALOWITCH, J.; ANASTASIOU, C.; MEHTA, S. V.; JAIN, L. K.; AGLIETTI, V.; JINDAL, D.; CHEN, P. et al. Big-bench extra hard. *arXiv preprint arXiv:2502.19187*, 2025.

KIM, M. J.; HAIZAN, I.; AHN, M. J.; PARK, D.-H.; CHOI, J.-H. Recent advances in lateral flow assays for viral protein detection with nanomaterial-based optical sensors. *Biosensors*, MDPI, v. 14, n. 4, p. 197, 2024.

KUMAR, R.; YADAV, S.; KUMAR, M.; KUMAR, J.; KUMAR, M. Artificial intelligence: new technology to improve indian agriculture. *International Journal of Chemical Studies*, v. 8, n. 2, p. 2999–3005, 2020.

KUMAR, S. S.; KHAN, A. K. M. A.; BANDAY, I. A.; GADA, M.; SHANBHAG, V. V. Overcoming llm challenges using rag-driven precision in coffee leaf disease remediation. In: IEEE. *2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*. [S.l.], 2024. p. 1–6.

KUMARAN, S.; SANJAY, P.; SANTHIYA, B. Improving agricultural productivity with cnn-based plant disease recognition system. In: IEEE. *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*. [S.l.], 2024. p. 289–294.

LANGCHAIN. *LangChain Introduction*. 2023. Https://docs.langchain.com.

LANGGRAPH. –. [Online; accessed 2025-10-24]. Available at: <https://www.langchain.com/langgraph>.

LANGGRAPH. *LangGraph Documentation*. 2024. Https://docs.langgraph.com.

LEWIS, P.; PEREZ, E.; PIKTUS, A.; PETRONI, F.; KARPUKHIN, V.; GOYAL, N.; KÜTTLER, H.; LEWIS, M.; YIH, W.-t.; ROCKTÄSCHEL, T. et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, v. 33, p. 9459–9474, 2020.

LEWIS, P.; PEREZ, E.; PIKTUS, A.; PETRONI, F.; KARPUKHIN, V.; GOYAL, N.; KULKARNI, H.; RIEDEL, S.; ZETTLEMOYER, L.; KIELA, D. Retrieval-augmented generation for knowledge-intensive nlp tasks. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2020. v. 33, p. 9459–9474.

LI, Y.; ZHAO, Z.; ZHANG, Q.; YU, L. Attention-based deep learning for plant disease detection and classification. *Computers and Electronics in Agriculture*, v. 187, p. 106285, 2021.

LIANG, J.; JIANG, W. A resnet50-dpa model for tomato leaf disease identification. *Frontiers in Plant Science*, Frontiers Media SA, v. 14, p. 1258658, 2023. Accuracy up to 97.60% on PlantVillage dataset.

LIANG, P.; BOMMASANI, R.; LEE, T.; TSIPRAS, D.; SOYLU, D.; YASUNAGA, M.; ZHANG, Y.; NARAYANAN, D.; WU, Y.; KUMAR, A. et al. Holistic evaluation of language models. *arXiv preprint arXiv:2211.09110*, 2022.

LIU, J.; WANG, X. Plant diseases and pests detection based on deep learning: a review. *Plant methods*, Springer, v. 17, n. 1, p. 22, 2021.

LIU, S. Y. Artificial intelligence (ai) in agriculture. *IT professional*, IEEE, v. 22, n. 3, p. 14–15, 2020.

MCP. –. [Online; accessed 2025-10-23]. Available at: <https://modelcontextprotocol.io/docs/getting-started/intro>.

MOHAMMADJAFARI, M. et al. Survey on text-to-sql using large language models. *arXiv preprint arXiv:2401.01467*, 2024.

MOHANTY, S. P.; HUGHES, D. P.; SALATHé, M. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, v. 7, p. 1419, 2016.

NÉMETH, M. Z.; KOVÁCS, G. M. A comprehensive guide to loop-mediated isothermal amplification, an emerging diagnostic tool for plant pathogenic fungi. *Frontiers in Plant Science*, Frontiers Media SA, v. 16, p. 1568657, 2025.

NEMMOUR, H.; MEZENNER, A.; ARAB, N.; LAKEHAL, M. R.; KHELFA, Z.; CHIBANI, Y. Deep learning ensemble based on transformer and cnn features for tomato leaf disease classification. In: SPIE. *Fifth Symposium on Pattern Recognition and Applications (SPRA 2024)*. [S.l.], 2025. v. 13540, p. 61–66.

PASCANU, R.; MIKOLOV, T.; BENGIO, Y. On the difficulty of training recurrent neural networks. In: PMLR. *International conference on machine learning*. [S.l.], 2013. p. 1310–1318.

PAWAR, A.; AMRUTKAR, R.; SOMANI, S.; MUNDADA, G. Plant leaf disease detection using cnn. In: AIP PUBLISHING LLC. *AIP Conference Proceedings*. [S.l.], 2024. v. 3156, n. 1, p. 070005.

PHAN, L.; GATTI, A.; HAN, Z.; LI, N.; HU, J.; ZHANG, H.; ZHANG, C. B. C.; SHAABAN, M.; LING, J.; SHI, S. et al. Humanity's last exam. *arXiv preprint arXiv:2501.14249*, 2025.

PHIDATA. –. [Online; accessed 2025-10-24]. Available at: <https://docs.phidata.com/introduction>.

PURANIK, S. S.; HANAMAKKANAVAR, S. R.; BIDARGADDI, A. P.; BALLUR, V. V.; JOSHI, P. T.; SM, M.; KULKARNI, U. Mobilenetv3 for mango leaf disease detection: An efficient deep learning approach for precision agriculture. In: IEEE. *2024 5th International Conference for Emerging Technology (INCET)*. [S.l.], 2024. p. 1–7.

QIN, Y.-M.; TU, Y.-H.; LI, T.; NI, Y.; WANG, R.-F.; WANG, H. Deep learning for sustainable agriculture: A systematic review on applications in lettuce cultivation. *Sustainability*, v. 17, n. 7, 2025. ISSN 2071-1050. Available at: <https://www.mdpi.com/2071-1050/17/7/3190>.

RAY, P. P. A survey on model context protocol: Architecture, state-of-the-art, challenges and future directions. *Authorea Preprints*, Authorea, 2025.

REZAYI, S.; LIU, Z.; WU, Z.; DHAKAL, C.; GE, B.; DAI, H.; MAI, G.; LIU, N.; ZHEN, C.; LIU, T.; LI, S. *Exploring New Frontiers in Agricultural NLP: Investigating the Potential of Large Language Models for Food Applications*. 2023. Available at: <https://arxiv.org/abs/2306.11892>.

REZAYI, S.; LIU, Z.; WU, Z.; DHAKAL, C.; GE, B.; DAI, H.; MAI, G.; LIU, N.; ZHEN, C.; LIU, T.; LI, S. Exploring New Frontiers in Agricultural NLP: Investigating the Potential of Large Language Models for Food Applications . *IEEE Transactions on Big Data*, IEEE Computer Society, Los Alamitos, CA, USA, v. 11, n. 03, p. 1235–1246, Jun. 2025. ISSN 2332-7790. Available at: <https://doi.ieeecomputersociety.org/10.1109/TBDATA.2024.3442542>.

RICHARDS, T. B. *Auto-GPT: An Autonomous GPT-4 Experiment*. 2023. Https://github.com/Torantulino/Auto-GPT.

RILEY, M. B.; WILLIAMSON, M. R.; MALOY, O. Plant disease diagnosis. *The plant health instructor*, The American Phytopathological Society, v. 10, p. 49, 2002.

ROUMELIOTIS, K. I.; SAPKOTA, R.; KARKEE, M.; TSELIKAS, N. D.; NASIOPOULOS, D. K. Plant disease detection through multimodal large language models and convolutional neural networks. *arXiv preprint arXiv:2504.20419*, 2025.

SALLAM, M.; BEAINI, C.; MIJWIL, M. M.; SALLAM, M. Microsoft ai diagnostic orchestrator (mai-dxo) promises cost savings in diagnosis—but at what cost to the medical profession? *Journal of Studies in Medical Sciences; Vol*, v. 1, n. 1, 2025.

SAPKOTA, R.; QURESHI, R.; HASSAN, S. Z.; SHUTSKE, J.; SHOMAN, M.; SAJJAD, M.; DHAREJO, F. A.; PAUDEL, A.; LI, J.; MENG, Z. et al. Multi-modal llms in agriculture: A comprehensive review. *Authorea Preprints*, Authorea, 2024.

SHI, W.; TANG, J.; LI, P.; LIU, P.; AL. et. A survey on retrieval-augmented generation. *arXiv preprint arXiv:2307.06958*, 2023.

SHOJAEE, P.; MIRZADEH, I.; ALIZADEH, K.; HORTON, M.; BENGIO, S.; FARAJTABAR, M. *The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity*. 2025. Available at: <https://ml-site.cdn-apple.com/papers/the-illusion-of-thinking.pdf>.

SINGHAL, K.; AZIZI, S.; TU, T.; MAHDAVI, S. S.; WEI, J.; CHUNG, H. W.; SCALES, N.; TANWANI, A.; COLE-LEWIS, H.; PFOHL, S. et al. Large language models encode clinical knowledge. *Nature*, Nature Publishing Group, v. 620, n. 7972, p. 172–180, 2023.

SLADOJEVIC, S.; ARSENOVIC, M.; ANDERLA, A.; CULIBRK, D.; STEFANOVIC, D. Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*, v. 2016, p. 3289801, 2016.

SRIVASTAVA, A.; RASTOGI, A.; RAO, A.; SHOEB, A. A.; ABID, A.; FISCH, A.; BROWN, A. R.; SANTORO, A.; GUPTA, A.; GARRIGA-ALONSO, A. et al. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *Transactions on machine learning research*, 2023.

SUN, Z.; ZHANG, Y.; TANG, J.; LIU, P.; AL. et. Reasoning over retrieval-augmented generation: Evaluating the factual consistency of language models with retrieved evidence. *arXiv preprint arXiv:2305.13250*, 2023.

SUNIL, C.; JAIDHAR, C.; PATIL, N. Systematic study on deep learning-based plant disease detection or classification. *Artificial Intelligence Review*, Springer, v. 56, n. 12, p. 14955–15052, 2023.

TEAM, O. I. *Open Interpreter*. 2023. Https://github.com/KillianLucas/open-interpreter.

TENG, Y. et al. Ai-driven agriculture: Integrating large language models with multimodal sensing for precision farming. *AI in Agriculture*, v. 9, p. 20–34, 2023.

THAKUR, P. S.; KHANNA, P.; SHEOREY, T.; OJHA, A. Explainable vision transformer enabled convolutional neural network for plant disease identification: Plantxvit. *arXiv preprint arXiv:2207.07919*, 2022.

TU, T.; SCHAEKERMANN, M.; PALEPU, A.; SAAB, K.; FREYBERG, J.; TANNO, R.; WANG, A.; LI, B.; AMIN, M.; CHENG, Y. et al. Towards conversational diagnostic artificial intelligence. *Nature*, Nature Publishing Group UK London, p. 1–9, 2025.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. *Advances in neural information processing systems*, v. 30, 2017.

VOORHEES, E. M.; TICE, D. M. The trec-8 question answering track report. In: *Text REtrieval Conference (TREC)*. [S.l.: s.n.], 1999.

WALEED, M.; UM, T.-W.; KAMAL, T.; KHAN, A.; IQBAL, A. Determining the precise work area of agriculture machinery using internet of things and artificial intelligence. *Applied Sciences*, MDPI, v. 10, n. 10, p. 3365, 2020.

WALSH, J. J.; MANGINA, E.; NEGRÃO, S. Advancements in imaging sensors and ai for plant stress detection: A systematic literature review. *Plant Phenomics*, AAAS, v. 6, p. 0153, 2024.

XIE, J.; CHEN, Z.; ZHANG, R.; WAN, X.; LI, G. Large multimodal agents: A survey. *arXiv preprint arXiv:2402.15116*, 2024.

YAN, F.; MAO, H.; JI, C. C.-J.; ZHANG, T.; PATIL, S. G.; STOICA, I.; GONZALEZ, J. E. Berkeley function calling leaderboard. In: . [S.l.: s.n.], 2024.

YING, J.; CHEN, Z.; WANG, Z.; JIANG, W.; WANG, C.; YUAN, Z.; SU, H.; KONG, H.; YANG, F.; DONG, N. *SeedBench: A Multi-task Benchmark for Evaluating Large Language Models in Seed Science*. 2025. Available at: <https://arxiv.org/abs/2505.13220>.

ZAHRT, D. *AI Wrkflow Automation Platform & Tools - n8n*. 2023. [Online; accessed 2025-10-23]. Available at: <https://n8n.io/>.

ZHANG, P.; GUO, Z.; ULLAH, S.; MELAGRAKI, G.; AFANTITIS, A.; LYNCH, I. Nanotechnology and artificial intelligence to enable sustainable and precision agriculture. *Nature Plants*, Nature Publishing Group UK London, v. 7, n. 7, p. 864–876, 2021.

ZHANG, X.; DU, C.; PANG, T.; LIU, Q.; GAO, W.; LIN, M. Chain of preference optimization: Improving chain-of-thought reasoning in llms. *Advances in Neural Information Processing Systems*, v. 37, p. 333–356, 2024.

ZHENG, C.; ZHANG, Z.; ZHANG, B.; LIN, R.; LU, K.; YU, B.; LIU, D.; ZHOU, J.; LIN, J. Processbench: Identifying process errors in mathematical reasoning. *arXiv preprint arXiv:2412.06559*, 2024.

ZHENG, L.; CHIANG, W.-L.; SHENG, Y.; ZHUANG, S.; WU, Z.; ZHUANG, Y.; LIN, Z.; LI, Z.; LI, D.; XING, E. et al. Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in neural information processing systems*, v. 36, p. 46595–46623, 2023.

ZHUGE, M.; ZHAO, C.; ASHLEY, D.; WANG, W.; KHIZBULLIN, D.; XIONG, Y.; LIU, Z.; CHANG, E.; KRISHNAMOORTHI, R.; TIAN, Y. et al. Agent-as-a-judge: Evaluate agents with agents. *arXiv preprint arXiv:2410.10934*, 2024.

APÊNDICE A – BOOKS AND REFERENCE MATERIALS

This appendix lists books and technical manuals that were consulted and used as knowledge sources for the Retrieval-Augmented Generation (RAG) component of this work.

BOOKS

Métodos em Fitopatologia - Acelino Couto Alfenas; Reginaldo Gonçalves Mafia. Editora UFV, 2nd ed., 2016. 516 pp. ISBN: 978-85-7269-559-6.

Manual de Fitopatologia: princípios e conceitos - Lilian Amorim; Jorge A. M. Rezende; Armando Bergamin Filho. 5th ed., 2018. 573 pp. ISBN: 978-85-318-0056-6.

Westcott's Plant Disease Handbook - R. K. Horst. Springer Reference, 8th ed., 2013. ISBN: 978-94-007-2140-1 (print). DOI: 10.1007/978-94-007-2141-8.

TECHNICAL MANUALS AND GUIDES

Hortaliças não-convencionais (tradicionais) - Ministério da Agricultura, Pecuária e Abastecimento (MAPA), Secretaria de Desenvolvimento Agropecuário e Cooperativismo. Brasília: MAPA/ACS, 2010.

Manual de Hortaliças Não-Convencionais - Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Brasília, 2010.

Manual de identificação e manejo de plantas daninhas em cultivos de cana-de-açúcar - Alexandre Magno Brighenti. Juiz de Fora: Embrapa Gado de Leite, 2010. 112 pp. ISBN: 978-85-7835-018-5.

Guia de diagnose para aulas práticas de fitopatologia: LFN 0424 – Fitopatologia - M. P. Gonçalves; A. L. T. Simões; R. F. dos Santos; S. de A. Lourenço; L. Amorim. 2nd ed., revised and expanded. Piracicaba: USP/ESALQ/LFN, 2022. 121 pp. ISBN: 978-65-87391-32-8. DOI: 10.11606/9786587391328.

Guia Prático de Plantas de Cobertura: aspectos fitotécnicos e impactos sobre a saúde do solo - Martha Lustosa Carvalho et al.; organized by Maurício Roberto Cherubin. Piracicaba: ESALQ-USP, 2022. 126 pp. ISBN: 978-65-89722-15-1. DOI: 10.11606/9786589722151.

APÊNDICE B – AGENT TEAMS PROMPTS

This appendix provides the English versions of the multi-agent prompt templates used in the system. These prompts were used as configuration and guidance for agents and were part of the knowledge context for the RAG-enabled assistants.

SUMMARY OF AGENT TEAMS

- **Agronomy Team** (supervisor + specialists)
  - *Soil Specialist*: collects soil context (type, pH, fertility, management) and recommends sustainable practices.
  - *Meteorologist*: assesses local climate (history/forecast) and impacts on crops; suggests weather-adapted practices.
  - *Crop Science Specialist*: reviews crop management (planting, fertilization, irrigation, pest/disease issues) and proposes improvements.
- **Botany Team** (supervisor + specialists)
  - *Taxonomist*: guides plant identification via morphology and distribution, requesting images and traits.
  - *Ecologist*: analyzes habitat and plant–environment interactions (abiotic/biotic factors).
- **Phytopathology Team** (supervisor + specialists)
  - *Mycologist*: triages and diagnoses fungal diseases; provides practical management.
  - *Virologist*: triages and diagnoses viral diseases; emphasizes vector control when applicable.
  - *Bacteriologist*: triages and diagnoses bacterial diseases; stresses preventive/sanitary measures.
- **Supervision Routing Flow**

  - Supervisor collects essential context (plant, symptoms, location, history, management, images).
  - Routes case by symptom patterns to the appropriate specialist; requests detailed, actionable recommendations.
  - Ensures responses are clear, locally applicable, and source-cited.

- **Knowledge Sources and RAG**
  - Specialists *must* consult the CultivAI Phytopathology knowledge base first; web search (Tavily) is secondary.
  - Prompts and references listed in the annexes served as knowledge inputs for the RAG pipeline.

PROMPTS (TEXT)

Soil Specialist Role: provide soil-based, sustainable land management guidance.

Objectives: assess soil physical, chemical, and biological properties; tailor recommendations to crop and local context.

Inputs to collect: soil type/classification, texture/structure, pH, EC/salinity, key nutrients (N, P, K, Ca, Mg, S, micronutrients), organic matter, CEC, drainage/compaction/erosion, recent and historical management (tillage, residues, rotations), fertilization and irrigation regimes, cropping system, observed constraints (toxicity, acidity, sodicity), lab analyses (attach values/units/dates).

Policy: consult the CultivAI knowledge base first (RAG), then web retrieval only to clarify specifics; cite sources.

Output: concise diagnosis of soil constraints, prioritized recommendations (amendments, fertilization plans, liming/gypsum, irrigation scheduling, cover crops, residue/tillage, erosion control), and a simple monitoring plan (what to measure and when).

Constraints: be practical, locally feasible, and specify rates/units/timing; highlight uncertainties and data gaps to confirm.

Meteorologist Role: translate climate conditions into crop-relevant actions. Inputs to collect: location (lat/long or municipality), elevation, recent weather (rain, temperature, humidity, wind), extremes (drought, flood, frost, heat), seasonal/weekly forecasts, ENSO or

regional outlooks. Policy: RAG with CultivAI first; web forecasts only as needed; cite forecast sources and timestamps. Output: risk assessment for the next 1–8 weeks (heat/frost, water stress, disease-conducive weather), recommended practices (irrigation scheduling, frost/heat mitigation, planting/harvest timing), and a short contingency plan for extremes. Constraints: align with crop phenology; quantify thresholds where possible.

Crop Science Specialist Role: optimize crop management for yield/quality and sustainability. Inputs to collect: crop(s)/varieties, planting dates/density, fertilization and irrigation practices, pest/disease issues observed, weed pressure, pruning/trellising (if any), growth stage, soil test summaries, history of yields and management. Policy: RAG with CultivAI first; web as secondary; cite. Output: prioritized actions (nutrient plan with rates/timing, irrigation targets, canopy/spacing adjustments, IPM actions, harvest/post-harvest guidance), plus expected outcomes and monitoring KPIs. Constraints: avoid conflicting actions with other specialists; reference local regulations when relevant.

Agronomy Supervisor Role: orchestrate Soil, Meteorology, and Crop Science. Steps: 1) gather essentials (crop, location, soil tests, weather, management, phenology, constraints, images); 2) route to the right specialist(s) with a clear subtask and needed outputs; 3) collect and reconcile recommendations; 4) ensure actions are specific (rates/units/timing), feasible, and non-contradictory; 5) deliver a unified plan with rationale and citations. Policy: enforce RAG order (CultivAI first); request missing info explicitly.

Taxonomist Role: guide plant identification. Inputs to collect: common/scientific names considered, location and habitat, morphological traits (habit, bark, leaves, flowers, fruits, seeds), phenology, images. Policy: use keys and regional floras from CultivAI knowledge first; web is secondary; cite keys used. Output: best-match taxon with rank, differential diagnosis vs. close taxa, confidence level, and references; request missing traits if uncertainty remains.

Ecologist Role: summarize plant–environment interactions. Inputs: location, habitat (climate, soil, vegetation type), biotic interactions (pollinators, dispersers, pests), disturbance regime, images if available. Output: concise ecology profile (niche, tolerances, community role), implications for cultivation or conservation, and key references. Constraints: be specific to the reported location/biome.

Botany Supervisor Role: coordinate Taxonomist and Ecologist. Steps: triage the request, ensure adequate morphological/location data, route to Taxonomist for identification and to

Ecologist for habitat/interaction summary, reconcile outputs, and present a coherent identification + ecology note with citations and uncertainty statements.

Mycologist Role: diagnose fungal diseases (e.g., rusts, mildews, anthracnose, rots). Inputs to collect: host plant and cultivar, location/season, symptom description and distribution, onset/progression, environmental conditions (humidity, temperature), recent treatments and cultural practices, images. Policy: consult CultivAI phytopathology knowledge first (RAG), then web if needed; cite sources, prefer CultivAI. Output: likely pathogen(s) with justification mapping symptoms to etiology, differential diagnosis, risk factors, management plan (cultural, biological, and chemical options with actives/rates/intervals where permissible), pre- and post-harvest notes, and preventive measures. Constraints: align with local registrations; clearly state when lab confirmation is recommended.

Virologist Role: diagnose viral diseases (mosaic, yellowing, dwarfing, deformation). Inputs: host/cultivar, vectors present/suspected, spatial pattern (systemic vs. localized), onset/progression, weather favoring vectors, prior controls. Policy: RAG order (CultivAI first); cite. Output: likely virus or virus group, transmission pathways, vector-focused IPM (monitoring, cultural, biological, chemical), sanitation/seed/cutting health recommendations, and guidance on confirmation testing (ELISA, PCR). Constraints: avoid speculative chemical advice; emphasize certified material and vector barriers.

Bacteriologist Role: diagnose bacterial diseases (e.g., Pseudomonas, Xanthomonas, Ralstonia). Inputs: host/cultivar, symptoms (water-soaked spots, exudates, cankers, wilt), humidity/rain influence, spread pattern, recent injuries or pruning, past controls. Policy: RAG with CultivAI first; cite. Output: likely bacterium or complex, sanitation and pruning hygiene, copper/bactericide considerations where allowed, irrigation/drainage adjustments, cultivar/rootstock notes, and lab confirmation guidance. Constraints: stress preventive/sanitary measures and resistance management.

Phytopathology Supervisor Role: triage plant problems and enforce diagnostic flow. Steps: 1) gather essentials (plant, symptoms, location, history, management, images); 2) route to Mycologist/Virologist/Bacteriologist according to symptom patterns; 3) ensure outputs specify likely pathogen, explain symptom links, and provide specific management with rates/intervals where relevant; 4) check internal consistency and add preventive measures and scouting schedule; 5) deliver a concise, cited summary. Policy: CultivAI first for RAG; web second.

APÊNDICE C – RAG Q&A DATASET

This appendix lists the first ten user questions (out of hundred) used for validation with their full reference answers and full reference contexts used by the RAG component. Reference contexts preserve source headings and extraction provenance (e.g., *p. X*, *Extraction hop: N*) generated by the RAG pipeline, for transparency.

- **Q1:** What are the symptoms of Phytophthora in avocados?

  Reference: *The symptoms of Phytophthora in avocados include general yellowing of the leaves, resembling nitrogen deficiency, followed by leaf drop and exposure of branches. There is also drying of the branch tips. Fruits rarely show symptoms, but there can be a sudden increase in the production of smaller fruits before the plant dies. The roots show discoloration and necrosis symptoms, with the fine roots almost completely destroyed. Bark cracking near the plant's collar, associated with gum exudation, can also be observed. Tissues just below the cracked bark show brown coloration and necrosis. Generally, the disease is only noticed at a very advanced stage, making control difficult and often leading to the plant's death.*

  Reference context:

  > GOMOSE - Phytophthora cinnamomi Rands
  >
  > **Sintomas:** A gomose ou podridão de raízes do abacateiro é uma das principais doenças da cultura tanto em viveiro como em campo. Sintomas desta doença são muito semelhantes aos da gomose dos citros, iniciando-se com amarelecimento generalizado das folhas, lembrando deficiência de nitrogênio. A seguir, ocorre queda das folhas e exposição dos ramos. Observa-se também seca de ramos do ponteiro. Frutos raramente apresentam sintomas da doença. É comum ocorrer, no entanto, um repentino aumento na produção de frutos menores na fase que antecede a morte das plantas. As raízes exibem descoloração e sintomas de necrose, e as radicelas ficam quase que totalmente destruídas. Fendilhamento da casca, na região próxima ao colo da planta, pode também ser observado, associado à exsudação de goma. Tecidos localizados logo abaixo da casca fendilhada apresentam coloração marrom e necrose. De um modo geral, a doença somente é percebida em estádio muito avançado, quando torna-se muito dificil seu controle, culminando com a morte da planta.

Etiologia: O fungo P. cinnamomi pertence à subdivisão Mastigomycotina e classe Oomycetes, apresentando hifa não-septada. O patógeno produz esporos assexuais, os zoósporos, que são liberados na presença de água e infectam o hospedeiro. Como estrutura de reprodução sexuada, o fungo produz oósporos, que apresentam paredes espessas e servem como estrutura de resistência. Esse patógeno tem boa capacidade saprofitica, podendo sobreviver por longos períodos desta forma. A sobrevivência do mesmo no solo e na ausência de plantas hospedeiras pode chegar até oito anos na forma de clamidósporo, e em raízes infectadas no mínimo 15 anos. O fungo necessita de água livre para que os zoósporos possam se locomover e infectar o hospedeiro. Portanto, a ocorrência da doença depende da presença de umidade elevada no solo, bem como de temperaturas entre 21 e 30°C. Temperaturas acima de 33°C inibem o desenvolvimento da doença completamente, enquanto que temperaturas entre 9 e 12°C reduzem muito sua incidência. Na literatura internacional são relatadas outras espécies de Phytophthora atacando o abacateiro, como P. cactovorum e P. citricola, que, normalmente não causam cancros, apenas podridões de raízes.

Controle: Medidas de controle incluem: a) uso de porta-enxertos tolerante ao fungo, como os PAGE 2 mexicanos Barr Duke, Duke, D9, Thomas, Toro Canyon, Borchard, Topa Topa e G-6; os guatemalenses G1033, Martin Grande (híbridos de R. americana com P. schiendeana Ness) G755a, G755b, G755c, UCR 2007, UCR 2008, UCR 2022, UCR 2023 e UCR 2053; e G-755 (P. schiedeana); b) aquisição ou produção de mudas de qualidade; c) remoção de restos de cultura tanto em viveiro como em campo; d) plantio de mudas em locais não encharcados; e) cuidados com o balanço nutricional. Níveis elevados de nitrogênio e pH e baixos de cálcio e fósforo aumentam a predisposição da planta à doença; f) evitar ferimentos nas raízes ou mesmo no tronco das árvores, pois constituem-se em vias de entrada do patógeno na planta; g) usar fungicidas quando a doença é constatada em seu início. Entre os fungicidas com possibilidade de uso temos: metalaxyl (aplicação via solo) e fosetyl.

- **Q2:** What are the recommended strategies for controlling Rosellinia necatrix in avocado cultivation, particularly concerning the susceptibility of Mexican varieties?

  Reference: *To control Rosellinia necatrix in avocado cultivation, it is advised to avoid*

planting in newly cleared areas or regions rich in organic matter. It is important to gather and burn crop residues and roots present in the soil, eliminate diseased plants and their root systems by burning them on-site if possible, and plow the soil. Additionally, planting in moist soils should be avoided, and care should be taken to prevent injuries to the plants, especially the roots, during cultivation operations. It is also crucial to use resistant rootstocks, as Mexican and Guatemalan varieties are very sensitive to the pathogen.

Reference context:

PODRIDÃO DE RAÍZES - Rosellinia necatrix Prill (Dematophora necatrix)

De maneira geral, a podridão de Rosellinia não tem grande importância econômica, sendo problema apenas em áreas isoladas. É uma doença típica de áreas recém-desbravadas, devido a alta capacidade saprofitica do patógeno.

Sintomas: Inicialmente observa-se murcha e sintomas que lembram deficiência nutricional, caracterizados por amarelecimento foliar. A doença manifesta-se de maneira lenta, levando alguns meses ou até anos para matar o hospedeiro. São comuns sintomas de murcha ou seca de folhas mais novas, ocasionando seca de ponteiros, que pode ocorrer por toda a planta ou apenas em algum lado da planta, correspondendo ao lado do sistema radicular afetado. Sintomas e sinais nas raízes caracterizam-se por podridão e coloração branca logo abaixo da casca.

Etiologia: Em geral o fungo ascomiceto R. necatrix, um parasita facultativo, é facilmente encontrado em restos de troncos, raízes mortas ou matéria orgânica devido à sua capacidade saprofitica. Em condições de elevada umidade, o patógeno pode formar cordões miceliais de coloração negra sobre as raízes ou sobre a matéria orgânica próxima à planta atacada. É comum também a presença de peritécios sobre raízes, quando o estado de podridão radicular mostra-se bem avançado.

Controle: Deve-se: evitar o plantio em áreas recém-desbravadas ou em regiões muito ricas em matéria orgânica; amontoar e queimar restos de cultura e raízes presentes no solo; eliminar plantas doentes e seus sistemas radiculares através da queima dos mesmos, se possível no próprio local, e alqueivar o solo; evitar o plantio em solos úmidos; evitar ferimentos nas plantas, principalmente nas raízes, durante as operações de cultivo; utilizar porta-

enxertos resistentes (as variedades mexicanas e guatemalenses são muito sensíveis ao patógeno).

- **Q3:** What does *D. ribis* do to avocado?

  Reference: *D. ribis is associated with symptoms similar to those caused by Dothiorella gregaria, which include canker and fruit rot in avocados. However, no studies have been conducted to verify the occurrence or evaluate the damage caused by D. ribis under the given conditions.*

  Reference context:

  CANCRO E PODRIDÃO DE FRUTO - Dothiorella gregaria Sacc.

  Sintomas: Podem ser observados tanto em ramos, tronco ou ainda em frutos, neste último caso sendo mais comuns em pós-colheita. Nos ramos e troncos, a doença manifesta-se através de fendilhamento e escamamento, sendo possível observar uma massa branca pulverulenta nos pontos de fendilhamento. Sintomas de cancro têm importância esporádica e ocorrem somente em algumas variedades. Locais afetados tendem a exibir descoloração e necrose dos vasos, interrompendo o fluxo normal da seiva, provocando a seca de ramos e podendo, inclusive, causar a morte da planta. O patógeno pode ocasionar danos no colo das plantas e, ocasionalmente, sintomas de seca dos ponteiros. Na superficie dos frutos ainda verdes, sintomas aparecem inicialmente como pequenas pontuações de coloração marrom ou púrpura. As lesões formadas aumentam de tamanho, até envolver o fruto completamente. O patógeno tende a invadir a polpa do abacate, ocasionando um escurecimento de tonalidade marrom e liberação de odor desagradável. Também pode ocorrer a queda prematura dos frutos, visto que o fungo pode infectar o pedúnculo dos mesmos.

  Etiologia: O agente causal tanto do cancro como das podridões de frutos é Dothiorella gregaria. Porém, na literatura encontramos D. ribis e D. aromatica associados a sintomas semelhantes. No entanto, até o momento não foram conduzidos trabalhos a fim de verificar a ocorrência ou não das demais espécies em nossas condições e avaliar os danos causados pela doença, principalmente em pós-colheita. O patógeno é beneficiado por alta umidade e presença de matéria orgânica, devido a sua capacidade saprofitica. Em geral, o

inóculo primário responsável pelas infecções nos frutos é oriundo de ramos secos.

Controle: Recomendam-se: eliminação de ramos secos ou debilitados, frutos com sintomas de podridões e árvores em produção com sintomas típicos da doença; plantio em locais bem drenados e sem excesso de matéria orgânica; aplicação regular de fungicidas cúpricos ou ditiocarbamatos após operações de poda; proteção de ferimentos com pasta cúprica; aplicação preventiva dos mesmos fungicidas, em 2 a 3 aplicações a partir de setembro, em áreas altamente afetadas; utilização de enxertia alta e de porta-enxertos resistentes e aplicação de fungicidas cúpricos na região de enxertia.

- **Q4:** What is the significance of Flórida in the context of avocado diseases?

  Reference: *Flórida is significant in the context of avocado diseases as it is where verrugose, or avocado scab, was first known in 1918. This disease is one of the main diseases affecting avocado trees, impacting the appearance and development of the fruit.*

  Reference context:

  VERRUGOSE - Sphaceloma perseae Jenkins

  A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

  Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

  Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O controle da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em condições de alta umidade, podem surgir alguns pontos de coloração acinzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas

- **Q5:** What is the resistance level of the Linda variety to *Cercospora purpurea*?

  Reference: *The Linda variety is considered to be moderately resistant to Cercospora purpurea.*

Reference context:

Cercospora purpurea e C. perseae.

A primeira é a única relatada até o momento no Estado de São Paulo. A incidência da doença inicia-se gradativamente na primeira metade do período chuvoso, atingindo um pico nos meses de PAGE 5 junho e julho. Nesse momento, inicia-se a queda das folhas. A sobrevivência do patógeno na cultura dá-se através das infecções foliares. Visto que a principal forma de disseminação do patógeno é por via aérea, a ocorrência da doença nos frutos é observada desde o início da frutificação. Controle Recomenda-se o uso de variedades resistentes, entre as quais as resistentes Collinson e Pollock (variedades antilhanas) e as medianamente resistentes Price, Simminds e Linda (variedades guatemalenses). Wagner é altamente suscetível (variedade guatemalense). O controle químico é complicado devido ao porte da planta e à inexistência de produtos de boa eficiência registrados para o uso na cultura. Porém, é possível a aplicação de cúpricos e ditiocarbamatos em casos onde a doença ocorre após a queda das folhas, pouco antes da florada do abacateiro, e logo após a queda de 2/3 das pétalas.

ANTRACNOSE Glomerella cingulata (Stonem) Spauld & Schrenk (Colletotrichum gloeosporioides (Penz.) Sacc.).

Sintomas: A antracnose afeta principalmente frutos, sendo possível encontrar o patógeno infectando folhas, flores e ramos, porém sem ocasionar danos à cultura. Sintomas em folhas são caracterizados por manchas necróticas de coloração escura, com bordos definidos e formato irregular. O patógeno pode ocorrer também nos ramos, causando necroses escuras e seca dos ramos e ponteiros, sendo este um sintoma de ocorrência rara. As flores podem ser facilmente afetadas pelo patógeno, ocorrendo seca ou abscisão das mesmas ou então serem infectadas através do botão floral, o que afetará o desenvolvimento do fruto, causando queda prematura e/ou podridão. Sintomas nos frutos são característicos, iniciando-se por pequenas pontuações de coloração marrom a preta, com formato circular e tamanho aproximado de 6-13 mm de diâmetro. As lesões tendem a evoluir atingindo parte do fruto ou necrosando-o completamente. As necroses ultrapassam a casca e alcançam a polpa do fruto. Uma vez dentro do fruto, o fungo causa um escurecimento da polpa de coloração marrom ou bege. É muito comum a ocorrência de frutos com po-

dridão no pedúnculo, a qual tem início nas infecções ocorridas nas flores ou em pós-colheita no ponto de cicatrização, caso ocorra a queda do pedúnculo. Em geral, este tipo de sintoma leva ao apodrecimento de todo o fruto, acarretando na planta a queda do mesmo. Podridões de frutos ocorrem em frutos maduros, sendo raros os efeitos em frutos verdes. A doença somente adquire importância em pomares mal tratados ou debilitados nutricionalmente.

Etiologia: O patógeno Colletotrichum gloeosporioides corresponde, na forma teleomórfica, a Glomerella cingulata. O fungo necessita de água livre para que ocorra a germinação e infecção, sendo a faixa ideal de temperatura para o crescimento 22-27℃. Permanece latente em frutos verdes, causando sintomas apenas após seu amadurecimento.

Controle: Deve ser realizado através de adubações e técnicas de manejo adequadas. Podas de limpeza e queima de material doente devem ser realizadas, no mínimo, anualmente. Ferimentos nos frutos devem ser evitados através de cuidados durante as operações de colheita e pós-colheita e controle de insetos.

- **Q6:** What are the symptoms and control methods for C. perseae in avocado cultivation?

  Reference: *C. perseae, associated with cercosporiosis, presents symptoms in avocado fruits as small, slightly depressed, irregular brown lesions with defined edges. Under high humidity, grayish spots may appear in the center due to pathogen sporulation. These lesions, approximately 3 to 6 mm in diameter, can lead to tissue fissures, allowing other pathogens to infect. Severe symptoms include fruit drop, causing significant production loss. On leaves, symptoms include angular brown or gray lesions with a chlorotic halo, 1 to 3 mm in diameter, visible on both sides, which may coalesce. Necrotic tissue in the lesion centers may fall out, causing leaf tearing. Control methods include using resistant varieties, particularly those from the Antillean group, which show high susceptibility to leaf verrugose but lower to fruit verrugose. Guatemalan varieties show high fruit susceptibility but low leaf susceptibility. Fungicidal applications, especially copper-based, are recommended, starting when 2/3 of the petals have fallen and continuing until fruits reach 5 cm in diameter. For leaves, control should occur during budding until they reach at least 3 cm in length.*

  Reference context:

  Extraction hop: 1

Cercospora purpurea e C. perseae.

A primeira é a única relatada até o momento no Estado de São Paulo. A incidência da doença inicia-se gradativamente na primeira metade do período chuvoso, atingindo um pico nos meses de PAGE 5 junho e julho. Nesse momento, inicia-se a queda das folhas. A sobrevivência do patógeno na cultura dá-se através das infecções foliares. Visto que a principal forma de disseminação do patógeno é por via aérea, a ocorrência da doença nos frutos é observada desde o início da frutificação. Controle Recomenda-se o uso de variedades resistentes, entre as quais as resistentes Collinson e Pollock (variedades antilhanas) e as medianamente resistentes Price, Simminds e Linda (variedades guatemalenses). Wagner é altamente suscetível (variedade guatemalense). O controle químico é complicado devido ao porte da planta e à inexistência de produtos de boa eficiência registrados para o uso na cultura. Porém, é possível a aplicação de cúpricos e ditiocarbamatos em casos onde a doença ocorre após a queda das folhas, pouco antes da florada do abacateiro, e logo após a queda de 2/3 das pétalas.

ANTRACNOSE Glomerella cingulata (Stonem) Spauld & Schrenk (Colletotrichum gloeosporioides (Penz.) Sacc.).

Sintomas: A antracnose afeta principalmente frutos, sendo possível encontrar o patógeno infectando folhas, flores e ramos, porém sem ocasionar danos à cultura. Sintomas em folhas são caracterizados por manchas necróticas de coloração escura, com bordos definidos e formato irregular. O patógeno pode ocorrer também nos ramos, causando necroses escuras e seca dos ramos e ponteiros, sendo este um sintoma de ocorrência rara. As flores podem ser facilmente afetadas pelo patógeno, ocorrendo seca ou abscisão das mesmas ou então serem infectadas através do botão floral, o que afetará o desenvolvimento do fruto, causando queda prematura e/ou podridão. Sintomas nos frutos são característicos, iniciando-se por pequenas pontuações de coloração marrom a preta, com formato circular e tamanho aproximado de 6-13 mm de diâmetro. As lesões tendem a evoluir atingindo parte do fruto ou necrosando-o completamente. As necroses ultrapassam a casca e alcançam a polpa do fruto. Uma vez dentro do fruto, o fungo causa um escurecimento da polpa de coloração marrom ou bege. É muito comum a ocorrência de frutos com podridão no pedúnculo, a qual tem início nas infecções ocorridas nas flores ou

em pós-colheita no ponto de cicatrização, caso ocorra a queda do pedúnculo. Em geral, este tipo de sintoma leva ao apodrecimento de todo o fruto, acarretando na planta a queda do mesmo. Podridões de frutos ocorrem em frutos maduros, sendo raros os efeitos em frutos verdes. A doença somente adquire importância em pomares mal tratados ou debilitados nutricionalmente.

Etiologia: O patógeno Colletotrichum gloeosporioides corresponde, na forma teleomórfica, a Glomerella cingulata. O fungo necessita de água livre para que ocorra a germinação e infecção, sendo a faixa ideal de temperatura para o crescimento 22-27℃. Permanece latente em frutos verdes, causando sintomas apenas após seu amadurecimento.

Controle: Deve ser realizado através de adubações e técnicas de manejo adequadas. Podas de limpeza e queima de material doente devem ser realizadas, no mínimo, anualmente. Ferimentos nos frutos devem ser evitados através de cuidados durante as operações de colheita e pós-colheita e controle de insetos.

Extraction hop: 2

VERRUGOSE - Sphaceloma perseae Jenkins

A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e

desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O controle da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em condições de alta umidade, podem surgir alguns pontos de coloração acinzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas

- **Q7:** What are the symptoms and control measures for C. perseae affecting avocado trees, and how does it compare to other diseases like verrugose and anthracnose?

Reference: *C. perseae, associated with cercosporiosis, presents symptoms on avocado fruits as small, slightly depressed, irregular brown lesions with defined edges. Under high humidity, grayish points may appear in the center due to pathogen sporulation. These lesions can lead to tissue fissures, allowing secondary infections, and cause significant fruit drop, leading to production loss. On leaves, symptoms include angular brown or gray lesions with a chlorotic halo, visible on both sides, which may coalesce and cause leaf tearing. Control involves using resistant varieties and applying copper-based fungicides. In comparison, verrugose, caused by Sphaceloma perseae, primarily affects fruits with warty, brown eruptions that coalesce but do not penetrate the skin. It also affects leaves, causing chocolate-colored spots and potential deformation. Control includes using resistant varieties and copper fungicides. Anthracnose, caused by Colletotrichum gloeosporioides, affects fruits, leaves, flowers, and branches, causing necrotic spots and fruit rot. Control involves proper nutrition, pruning, and avoiding fruit injuries. Each disease requires specific management strategies to minimize impact on avocado cultivation.*

Reference context:

Extraction hop: 1

Cercospora purpurea e C. perseae.

A primeira é a única relatada até o momento no Estado de São Paulo. A incidência da doença inicia-se gradativamente na primeira metade do período chuvoso, atingindo um pico nos meses de PAGE 5 junho e julho. Nesse momento, inicia-se a queda das folhas. A sobrevivência do patógeno na cultura dá-se através das infecções foliares. Visto que a principal forma de disseminação do patógeno é por via aérea, a ocorrência da doença nos frutos é observada desde o início da frutificação.

Controle: Recomenda-se o uso de variedades resistentes, entre as quais as resistentes Collinson e Pollock (variedades antilhanas) e as medianamente resistentes Price, Simminds e Linda (variedades guatemalenses). Wagner é altamente suscetível (variedade guatemalense). O controle químico é complicado devido ao porte da planta e à inexistência de produtos de boa eficiência registrados para o uso na cultura. Porém, é possível a aplicação de cúpricos e ditiocarbamatos em casos onde a doença ocorre após a queda das folhas, pouco antes da florada do abacateiro, e logo após a queda de 2/3 das pétalas.

ANTRACNOSE Glomerella cingulata (Stonem) Spauld & Schrenk (Colletotrichum gloeosporioides (Penz.) Sacc.).

Sintomas: A antracnose afeta principalmente frutos, sendo possível encontrar o patógeno infectando folhas, flores e ramos, porém sem ocasionar danos à cultura. Sintomas em folhas são caracterizados por manchas necróticas de coloração escura, com bordos definidos e formato irregular. O patógeno pode ocorrer também nos ramos, causando necroses escuras e seca dos ramos e ponteiros, sendo este um sintoma de ocorrência rara. As flores podem ser facilmente afetadas pelo patógeno, ocorrendo seca ou abscisão das mesmas ou então serem infectadas através do botão floral, o que afetará o desenvolvimento do fruto, causando queda prematura e/ou podridão. Sintomas nos frutos são característicos, iniciando-se por pequenas pontuações de coloração marrom a preta, com formato circular e tamanho aproximado de 6-13 mm de diâmetro. As lesões tendem a evoluir atingindo parte do fruto ou necrosando-o completamente. As necroses ultrapassam a casca e alcançam a polpa do fruto. Uma vez dentro do fruto, o fungo causa um escurecimento da polpa de coloração marrom ou bege. É muito comum a ocorrência de frutos com podridão no pedúnculo, a qual tem início nas infecções ocorridas nas flores ou em pós-colheita no ponto de cicatrização, caso ocorra a queda do pedúnculo. Em geral, este tipo de sintoma leva ao apodrecimento de todo o fruto, acarretando na planta a queda do mesmo. Podridões de frutos ocorrem em frutos maduros, sendo raros os efeitos em frutos verdes. A doença somente adquire importância em pomares mal tratados ou debilitados nutricionalmente.

Etiologia: O patógeno Colletotrichum gloeosporioides corresponde, na forma teleomórfica, a Glomerella cingulata. O fungo necessita de água livre para que ocorra a germinação e infecção, sendo a faixa ideal de temperatura para o crescimento 22-27℃. Permanece latente em frutos verdes, causando sintomas apenas após seu amadurecimento.

Controle: Deve ser realizado através de adubações e técnicas de manejo adequadas. Podas de limpeza e queima de material doente devem ser realizadas, no mínimo, anualmente. Ferimentos nos frutos devem ser evitados através de cuidados durante as operações de colheita e pós-colheita e controle de insetos.

Extraction hop: 2

VERRUGOSE - Sphaceloma perseae Jenkins

A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O controle da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em condições de alta umidade, podem surgir alguns pontos de coloração ac-

inzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas

- **Q8:** What are the symptoms and control methods for S. perseae affecting avocado trees?

  Reference: *The symptoms of S. perseae, also known as verrugose or avocado scab, primarily appear on the fruits as small, eruptive, wart-like spots that are 5 to 6 mm in size and brown in color. These spots can rapidly increase in size and coalesce, but the infection never penetrates beyond the fruit's skin. On leaves, symptoms manifest as small, chocolate-colored spots, 1 to 2 mm in diameter, which are round on the leaf blade and slightly elongated on the veins, resembling scale insects. Severely affected leaves may deform, rupture, and have reduced photosynthetic area. The disease is caused by the fungus S. perseae, which attacks leaves up to 3 cm long and fruits less than 5 cm in diameter, thriving in high humidity conditions. Control methods include using resistant varieties, with Antillean group varieties being highly susceptible to leaf verrugose and less so to fruit, while Guatemalan group varieties show high susceptibility in fruits and low in leaves. Copper-based fungicides are recommended, starting when at least two-thirds of the petals have fallen and continuing until fruits reach 5 cm in diameter. For leaves, control should be applied during sprouting periods until they reach at least 3 cm in length. In nurseries, Guatemalan group varieties should receive bi-weekly copper fungicide applications.*

  Reference context:

  Extraction hop: 1

VERRUGOSE - Sphaceloma perseae Jenkins

A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O controle da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em

condições de alta umidade, podem surgir alguns pontos de coloração acinzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas.

- **Q9:** What are the symptoms and control measures for C. perseae affecting avocado trees, and how does it compare to other diseases like verrugose and cercosporiose?

  Reference: *C. perseae, associated with cercosporiose, causes small, slightly depressed, irregular brown lesions on avocado fruits, which can lead to fissures and secondary infections. On leaves, it creates angular brown or gray lesions with a chlorotic halo, which can coalesce and cause tearing. Control involves using resistant varieties and applying fungicides like cuprics and dithiocarbamates. Verrugose, caused by Sphaceloma perseae, results in warty, eruptive lesions on fruits and chocolate-colored spots on leaves, leading to deformation and reduced photosynthesis. Control includes using resistant varieties and fungicides. Both diseases thrive in high humidity, but verrugose primarily affects young fruits and leaves, while cercosporiose can cause significant fruit drop.*

  Reference context:

  Extraction hop: 1

  Cercospora purpurea e C. perseae.

  A primeira é a única relatada até o momento no Estado de São Paulo. A incidência da doença inicia-se gradativamente na primeira metade do período chuvoso, atingindo um pico nos meses de PAGE 5 junho e julho. Nesse mo-

mento, inicia-se a queda das folhas. A sobrevivência do patógeno na cultura dá-se através das infecções foliares. Visto que a principal forma de disseminação do patógeno é por via aérea, a ocorrência da doença nos frutos é observada desde o início da frutificação.

Controle: Recomenda-se o uso de variedades resistentes, entre as quais as resistentes Collinson e Pollock (variedades antilhanas) e as medianamente resistentes Price, Simminds e Linda (variedades guatemalenses). Wagner é altamente suscetível (variedade guatemalense). O controle químico é complicado devido ao porte da planta e à inexistência de produtos de boa eficiência registrados para o uso na cultura. Porém, é possível a aplicação de cúpricos e ditiocarbamatos em casos onde a doença ocorre após a queda das folhas, pouco antes da florada do abacateiro, e logo após a queda de 2/3 das pétalas.

ANTRACNOSE Glomerella cingulata (Stonem) Spauld & Schrenk (Colletotrichum gloeosporioides (Penz.) Sacc.).

Sintomas: A antracnose afeta principalmente frutos, sendo possível encontrar o patógeno infectando folhas, flores e ramos, porém sem ocasionar danos à cultura. Sintomas em folhas são caracterizados por manchas necróticas de coloração escura, com bordos definidos e formato irregular. O patógeno pode ocorrer também nos ramos, causando necroses escuras e seca dos ramos e ponteiros, sendo este um sintoma de ocorrência rara. As flores podem ser facilmente afetadas pelo patógeno, ocorrendo seca ou abscisão das mesmas ou então serem infectadas através do botão floral, o que afetará o desenvolvimento do fruto, causando queda prematura e/ou podridão. Sintomas nos frutos são característicos, iniciando-se por pequenas pontuações de coloração marrom a preta, com formato circular e tamanho aproximado de 6-13 mm de diâmetro. As lesões tendem a evoluir atingindo parte do fruto ou necrosando-o completamente. As necroses ultrapassam a casca e alcançam a polpa do fruto. Uma vez dentro do fruto, o fungo causa um escurecimento da polpa de coloração marrom ou bege. É muito comum a ocorrência de frutos com podridão no pedúnculo, a qual tem início nas infecções ocorridas nas flores ou em pós-colheita no ponto de cicatrização, caso ocorra a queda do pedúnculo. Em geral, este tipo de sintoma leva ao apodrecimento de todo o fruto, acarretando na planta a queda do mesmo. Podridões de frutos ocorrem em frutos maduros, sendo raros os efeitos em frutos verdes. A doença somente adquire

importância em pomares mal tratados ou debilitados nutricionalmente.

Etiologia: O patógeno Colletotrichum gloeosporioides corresponde, na forma teleomórfica, a Glomerella cingulata. O fungo necessita de água livre para que ocorra a germinação e infecção, sendo a faixa ideal de temperatura para o crescimento 22-27℃. Permanece latente em frutos verdes, causando sintomas apenas após seu amadurecimento.

Controle: Deve ser realizado através de adubações e técnicas de manejo adequadas. Podas de limpeza e queima de material doente devem ser realizadas, no mínimo, anualmente. Ferimentos nos frutos devem ser evitados através de cuidados durante as operações de colheita e pós-colheita e controle de insetos.

Extraction hop: 2

VERRUGOSE - Sphaceloma perseae Jenkins

A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua

vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O controle da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em condições de alta umidade, podem surgir alguns pontos de coloração acinzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas

- **Q10:** What are the simptoms and control methods for the disease verrugose in avocado trees?

  Reference: *The disease caused by S. perseae, known as verrugose or avocado scab, presents symptoms primarily on the fruits as small, eruptive, wart-like spots that are 5 to 6 mm in size and brown in color. These spots can rapidly increase in size and coalesce, but the infection does not penetrate beyond the fruit's skin. On leaves, symptoms appear*

as small, chocolate-colored spots, 1 to 2 mm in diameter, which are round on the leaf blade and slightly elongated on the veins. Severely affected leaves may deform, rupture, and have reduced photosynthetic area. The disease develops under high humidity conditions. Control methods include using resistant varieties, with Antillean group varieties being highly susceptible to leaf verrugose and less so to fruit, while Guatemalan group varieties show high susceptibility in fruits and low in leaves. Copper-based fungicides are recommended, starting when two-thirds of the petals have fallen and continuing until the fruits reach 5 cm in diameter. For leaves, control should be applied during sprouting periods until they reach at least 3 cm in length. In nurseries, Guatemalan group varieties should receive bi-weekly copper fungicide applications.

Reference context:

Extraction hop: 1

VERRUGOSE - Sphaceloma perseae Jenkins

A verrugose, ou sarna do abacateiro, conhecida desde 1918 na Flórida, foi encontrada no Brasil pela PAGE 4 primeira vez em 1938 em Limeira. É uma das principais doenças do abacateiro, visto que a mesma, além de depreciar a aparência do fruto, pode provocar também a queda de frutos jovens bem como o subdesenvolvimento em situações de alta severidade de doença.

Sintomas: São observados principalmente nos frutos, na forma de pequenas pontuações eruptivas, verrugosas, com tamanho de 5 a 6 mm de coloração marrom, que aumentam rapidamente e coalescem. A infecção nos frutos nunca ultrapassa a casca. A doença também pode ocasionar sintomas em folhas, na forma de pequenas pontuações de cor chocolate, com 1 a 2 mm de diâmetro, arredondadas quando localizadas no limbo foliar e ligeiramente alongadas quando nas nervuras, lembrando cochonilhas. Quando severamente atacadas, as folhas tendem a deformar e até mesmo sofrer rompimento do limbo foliar, além de redução da área fotossintética.

Etiologia: A doença é ocasionada pelo fungo S. perseae, que ataca folhas com no máximo 3 cm de comprimento e frutos com menos de 5 cm e desenvolve-se somente em condições de umidade elevada.

Controle: Recomenda-se a utilização de variedades resistentes. Variedades pertencentes ao grupo antilhano apresentam elevada suscetibilidade à verrugose das folhas e menor de fruto. Variedades do grupo guatemalense, por sua vez, apresentam elevada suscetibilidade nos frutos e baixa nas folhas. O con-

trole da doença pode também ser feito com a aplicação de fungicidas cúpricos. No caso dos frutos, deve-se iniciar o controle quando pelo menos 2/3 das pétalas caírem e mantê-lo até os frutos atingirem 5 cm de diâmetro. Para as folhas, o controle deve ser feito somente nos períodos de brotações até que as mesmas atinjam um mínimo de 3 cm de comprimento. Em viveiro de mudas, para variedades do grupo guatemalense, deve-se realizar aplicação quinzenal de fungicidas cúpricos.

CERCOSPORIOSE - Cercospora purpurea Cooke, C. perseae Ellis & Martin

Esta doença é muito importante nos cultivos de abacate da América Latina e Flórida.

Sintomas: Nos frutos são caracterizados por pequenas lesões, ligeiramente deprimidas e irregulares, de coloração marrom e bordos definidos. Em condições de alta umidade, podem surgir alguns pontos de coloração acinzentada no centro das lesões, os quais correspondem à esporulação do patógeno. Lesões nos frutos apresentam tamanhas aproximadas de 3 a 6 mm de diâmetro e, com o envelhecimento, tendem a provocar fissuras nos tecidos, possibilitando a infecção por outros patógenos. A queda de frutos é um dos sintomas mais severos da doença, podendo acarretar elevada perda na produção. Sintomas nas folhas caracterizam-se pela presença de lesões angulares de coloração marrom ou cinza, com halo clorótico. As lesões apresentam tamanho de 1 a 3 mm de diâmetro e são visíveis nas duas faces da folha, tendendo a coalescer. Tecidos necrosados no centro das lesões tendem a cair, facilitando o rasgamento do limbo foliar. As lesões podem ocorrer também no pedúnculo dos frutos, o que induz a queda dos mesmos. Essas lesões mostram-se muito semelhantes às do fruto, porém de coloração escura, formato circular e tamanho aproximado de 1 a 5 mm.

Etiologia: No Brasil foram encontradas 2 espécies de Cercospora associadas