



*Ontologia da Produção de Tabaco:  
enriquecendo o suporte à decisão baseado em  
LLM*

**Luis Felipe Medeiro Alves**

Dezembro / 2024

Dissertação de Mestrado em Ciência da  
Computação

# **Ontologia da Produção de Tabaco: enriquecendo o suporte à decisão baseado em LLM**

Esse documento corresponde à Dissertação apresentada à Banca Examinadora no curso de Mestrado em Ciência da Computação da UNIFACCAMP - Centro Universitario Campo Limpo Paulista.

Campo Limpo Paulista, 12 de dezembro de 2024.

Luis Felipe Medeiro Alves

Prof. Dr. Ferruccio de Franco Rosa (Orientador)

Ficha catalográfica elaborada pela  
Biblioteca Central da Unifaccamp

A48o

Alves, Luis Felipe Medeiro

Ontologia da produção de tabaco: enriquecendo o suporte à decisão baseado em LLM / Luis Felipe Medeiro Alves. Campo Limpo Paulista, SP: Unifaccamp, 2024.

58 p.: il.

Orientador: Prof. Dr. Ferrucio de Franco Rosa

Dissertação (Programa de Mestrado Profissional em Ciência da Computação) – Centro Universitário Campo Limpo Paulista – Unifaccamp.

1. Ontologia. 2. Ciência da Computação. 3. Agricultura. I. Rosa, Ferrucio de Franco. II. Centro Universitário Campo Limpo Paulista. III. Título.

CDD – 338.1

# Agradecimientos

An Gott.

# Resumo

A produção de tabaco (*Nicotiana tabacum*) é uma área com grande relevância econômica, mas ainda carece de soluções tecnológicas avançadas que integrem a otimização das práticas agrícolas com sistemas de inteligência artificial modernos. Esta dissertação apresenta a Ontologia da Produção de Tabaco (Ontology of Tobacco Production - OnTop), uma ontologia formal projetada para modelar o ciclo de vida da produção de tabaco com uma abordagem robusta para a integração com Modelos de Linguagem Extensos (LLMs). Baseada em princípios para o design de ontologias, OnTop se diferencia ao oferecer um framework claro e extensível que minimiza vieses de codificação, enquanto sustenta a interoperabilidade e inferência automatizada em sistemas computacionais. A dissertação aborda inicialmente os princípios de design de ontologias e sua aplicação na construção de OnTop, com o objetivo de enriquecer os modelos de linguagem e ampliar sua capacidade de fornecer recomendações precisas e contextualizadas no domínio agrícola. Em seguida, o trabalho apresenta uma revisão sistemática de ontologias aplicadas à inteligência artificial no setor agrícola, evidenciando a importância de abordagens computacionais específicas de domínio. A ontologia é descrita em detalhes, destacando suas classes e relações lógicas, e demonstrando como sua estrutura formal pode ser usada para melhorar a precisão e o contexto das respostas geradas por LLMs. A validação da OnTop, realizada por meio de sua aplicação em LLMs, evidencia o potencial das ontologias para aprimorar os sistemas de inteligência artificial na resolução de problemas complexos e na interoperabilidade semântica. Por fim, a conclusão discute as contribuições de OnTop para a ciência da computação, focando na formalização do conhecimento e na criação de bases sólidas para avanços futuros na integração entre ontologias e LLMs.

**Palavras-chaves:** Ontologia; Inteligência Artificial; Princípios de Gruber; Modelos de Linguagem Extensos; Formalização do Conhecimento; Interoperabilidade Semântica; Agricultura de Precisão; Integração Ontológica; Inferência Automatizada.

# Abstract

Tobacco (*Nicotiana tabacum*) production is a domain with high economic significance, yet it lacks advanced technological solutions that integrate optimized agricultural practices with modern artificial intelligence systems. This thesis presents the Ontology of Tobacco Production (OnTop), a formal ontology designed to model the tobacco production lifecycle with a robust approach for integration with Large Language Models (LLMs). Grounded in ontology design principles, OnTop provides a clear and extensible framework that minimizes encoding bias while supporting automated inference and semantic interoperability in computational systems. The thesis begins by discussing ontology design principles and their application in constructing OnTop, aimed at enriching language models and enhancing their ability to provide accurate, context-aware recommendations in the agricultural domain. It then presents a systematic review of ontologies applied to AI in agriculture, highlighting the need for domain-specific computational approaches. OnTop is described in detail, with a focus on its classes and logical relationships, demonstrating how its formal structure can be used to improve the accuracy and contextual relevance of LLM outputs. The validation of OnTop, through its application in LLMs, shows the potential of ontologies to enhance AI systems' capabilities in complex problem-solving and semantic interoperability. Finally, the conclusion discusses OnTop's contributions to computer science, focusing on its role in knowledge formalization and laying a solid foundation for future advancements in ontology-LLM integration.

**Keywords:** Ontology; Artificial Intelligence; Gruber's Principles; Large Language Models; Knowledge Formalization; Semantic Interoperability; Precision Agriculture; Ontological Integration; Automated Inference.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>11</b>
1.1	Questão de Pesquisa, Objetivos e Contribuições	12
1.2	Estrutura da Dissertação	13
<b>2</b>	<b>Referencial Teórico e Metodológico</b>	<b>14</b>
2.1	Conceituação de ontologias por Gruber	14
2.2	Síntese dos Fundamentos da Produção de Tabaco	17
<b>3</b>	<b>Revisão de Literatura e Trabalhos Relacionados</b>	<b>20</b>
3.1	Sistemas de Recomendação Baseados em GIS na Agricultura	20
3.1.1	Metodologia da Revisão Sistemática	20
3.1.2	Resultados e Discussão	21
3.2	Trabalhos Relacionados	24
<b>4</b>	<b>Ontologia da Produção de Tabaco</b>	<b>27</b>
4.1	Exemplos de Definições Formais	33
4.2	Avaliação de OnTop	37
<b>5</b>	<b>Aplicando OnTop em um Modelo de Linguagem Extenso (LLM)</b>	<b>45</b>
<b>6</b>	<b>Discussão sobre a Aplicação</b>	<b>48</b>
<b>7</b>	<b>Conclusão</b>	<b>51</b>
7.1	Resultados da Pesquisa	51
7.2	Limitações e Trabalhos Futuros	52
7.3	Considerações éticas	53
7.4	Resultados Técnico-científicos	54
	<b>Referências</b>	<b>55</b>

# Lista de tabelas

Tabela 2.1 – Conceitos paralelos entre ontologias e conjuntos . . . . .	14
Tabela 3.1 – Síntese comparativa dos estudos analisados . . . . .	22

# Lista de ilustrações

Figura 4.1 – Tela inicial do logiciário <i>Protégé</i> . . . . .	28
Figura 4.2 – Classes Principais de OnTop.owl . . . . .	28
Figura 4.3 – Principais classes e subclasses de OnTop . . . . .	31
Figura 4.4 – A descrição dos usos de uma classe no <i>Protégé</i> . . . . .	32
Figura 4.5 – Exemplo do uso de SWRL no <i>Protégé</i> no contexto de OnTop . . . . .	32
Figura 4.6 – Relação entre <b><i>TobaccoSeedbed</i></b> e <b><i>TobaccoSeed</i></b> . . . . .	35
Figura 4.7 – Classe <b><i>SoilCondition</i></b> , passível de expansão . . . . .	36
Figura 4.8 – Análise de OnTop por meio de OntoMetrics . . . . .	38
Figura 4.9 – Exemplo do nível de detalhamento da Classe <b><i>ManagementAction</i></b> . . . . .	39
Figura 4.10–Exemplos de propriedades de OnTop. . . . .	41
Figura 4.11–Indivíduos declarados em OnTop . . . . .	43
Figura 4.12–Exemplos de indivíduos relacionados a <b><i>TobaccoSeedbed</i></b> . . . . .	44
Figura 4.13–Exemplo de anotação na Classe <b><i>ClimateCondition</i></b> . . . . .	44

# Lista de Siglas

<b>AI</b>	<i>Artificial Intelligence</i> (Inteligência Artificial)
<b>DL</b>	<i>Description Logic</i> (Lógica Descritiva)
<b>GIS</b>	<i>Geographic Information System</i> (Sistema de Informação Geográfica)
<b>LLM</b>	<i>Large Language Model</i> (Modelo de Linguagem Extenso ou Modelo de Linguagem de Grande Escala ou Modelo de Linguagem Amplo)
<b>OnTop</b>	<i>Ontology of Tobacco Production</i> (Ontologia de Produção de Tabaco)
<b>OWL</b>	<i>Web Ontology Language</i> (Linguagem de Ontologia para a Web)
<b>RS</b>	<i>Recommender System</i> (Sistema de Recomendação)
<b>SWRL</b>	<i>Semantic Web Rule Language</i> (Linguagem de Regras da Web Semântica)
<b>UT-LLM</b>	<i>Under-Testing Large Language Model</i> (Modelo de Linguagem Extensa em Teste)
<b>W3C</b>	<i>World Wide Web Consortium</i>

# 1 Introdução

A produção de tabaco (*Nicotiana tabacum*) desempenha um papel econômico crucial em várias regiões do mundo, especialmente em países em desenvolvimento. No Brasil, por exemplo, a cultura do tabaco é de grande relevância, sendo responsável por empregar cerca de 600 mil pessoas e gera 2 bilhões de dólares de receita anual (AGROPECUÁRIA, 2020). Apesar de sua importância econômica, a produção de tabaco enfrenta desafios relacionados à otimização de práticas agrícolas, sustentabilidade e integração tecnológica.

A agricultura moderna tem se beneficiado de avanços em tecnologias de informação e comunicação, especialmente no que tange a Sistemas de Informação Geográfica (GIS) e sistemas de recomendação (RS) baseados em inteligência artificial (PIERCE *et al.*, 2007). No entanto, para cultivos de tabaco, os produtores ainda estão em processo de adoção de inovações, preenchendo, vagarosamente, essas lacunas na eficiência produtiva e na gestão sustentável dos recursos.

Uma ontologia formal para a produção de tabaco pode fornecer uma estrutura sistemática para representar e compartilhar conhecimento especializado, facilitando a comunicação entre diferentes sistemas e agentes (GRUBER, 1995). A utilização de ontologias permite a integração de dados heterogêneos, promove a interoperabilidade de sistemas e apoia a tomada de decisão baseada em conhecimento consolidado.

Além de sua aplicação prática no setor agrícola, este trabalho procura contribuir também para a ciência da computação, pois aborda desafios fundamentais relacionados à formalização do conhecimento, à interoperabilidade semântica e à integração de ontologias com sistemas de inteligência artificial avançados, como os Modelos de Linguagem Extensos (LLMs). A criação da OnTop exemplifica a aplicação de princípios computacionais na modelagem de domínios complexos, promovendo o uso de linguagens formais, como OWL, para representar informações de maneira lógica e estruturada.

Neste contexto, a proposta de uma Ontologia da Produção de Tabaco (*Ontology of Tobacco Production* - OnTop) surge como uma contribuição inovadora, buscando formalizar todo o ciclo da produção de tabaco e integrar fatores ambientais, agrônômicos e gerenciais. OnTop visa preencher a lacuna existente na representação formal do conhecimento relacionado à cultura do tabaco, oferecendo um framework extensível para apoiar futuras inovações tecnológicas na agricultura. A implementação de ontologias na agricultura não é um conceito novo; entretanto, a aplicação específica para a produção de tabaco é inédita.

## 1.1 Questão de Pesquisa, Objetivos e Contribuições

A estrutura desta pesquisa foi baseada no seguinte questionamento: *Como a construção de uma ontologia formal específica para a produção de tabaco pode otimizar práticas agrícolas e facilitar a organização e uso do conhecimento nessa área?*

O objetivo geral desta dissertação é desenvolver uma ontologia formal para a produção de tabaco que possibilite a representação estruturada e o compartilhamento de conhecimento no domínio, facilitando a interoperabilidade entre sistemas e a tomada de decisão informada.

Os objetivos específicos são:

1. Analisar os princípios de design de ontologias propostos por Gruber (GRUBER, 1995) e aplicá-los no desenvolvimento da OnTop.
2. Realizar uma revisão sistemática da literatura sobre sistemas de recomendação baseados em GIS na agricultura, identificando lacunas e oportunidades para a aplicação na cultura do tabaco.
3. Projetar e formalizar OnTop, definindo suas principais classes, relações e restrições lógicas, garantindo clareza, coerência e extensibilidade.
4. Validar a aplicação da OnTop integrando-a a um Modelo de Linguagem Extensa (LLM), avaliando sua capacidade de aprimorar respostas e recomendações no contexto agrícola.

As contribuições desta dissertação incluem:

- Proposição de uma primeira ontologia formal dedicada à produção de tabaco, preenchendo uma lacuna significativa na representação do conhecimento neste domínio.
- Exemplificação da aplicabilidade dos princípios de Gruber na construção de ontologias voltadas para a agricultura, oferecendo um modelo extensível para outras culturas.
- Integração da OnTop a sistemas baseados em IA, exemplificando como ontologias podem enriquecer modelos de linguagem e melhorar a qualidade de recomendações e respostas em contextos especializados.

## 1.2 Estrutura da Dissertação

Esta dissertação está organizada em sete capítulos, além desta introdução. A seguir, apresentamos uma visão geral de cada capítulo:

- **Capítulo 2 - Referencial Teórico:** Aborda os princípios de design de ontologias propostos por Gruber (GRUBER, 1995), estabelecendo a base teórica para o desenvolvimento da OnTop. São discutidos conceitos fundamentais e critérios para a construção de ontologias voltadas ao compartilhamento de conhecimento.
- **Capítulo 3 - Revisão da Literatura:** Apresenta uma revisão sistemática da literatura sobre sistemas de recomendação baseados em GIS na agricultura. Este capítulo identifica tendências, desafios e a relevância de ontologias neste contexto.
- **Capítulo 4 - Ontologia da Produção de Tabaco:** Descreve em detalhes a OnTop e são exploradas as principais classes, relações e a formalização lógica empregada na ontologia.
- **Capítulo 5 - Aplicando OnTop em uma LLM:** Explora-se a integração da OnTop a um Modelo de Linguagem Extensa, avaliando como a ontologia pode melhorar a capacidade de tais modelos em fornecer respostas precisas e contextualizadas no domínio da produção de tabaco.
- **Capítulo 6 - Discussão dos Resultados:** Analisa os resultados obtidos com a aplicação da OnTop na LLM, discutindo as implicações, benefícios e possíveis limitações identificadas durante a validação prática.
- **Capítulo 7 - Conclusão:** Apresenta as considerações finais, destacando as contribuições do trabalho, suas limitações e sugestões para pesquisas futuras.

Ao longo da dissertação, buscamos fornecer uma visão abrangente sobre a importância e o impacto da utilização de ontologias na agricultura, especificamente na produção de tabaco. Espera-se que este trabalho contribua para o avanço tecnológico no setor agrícola, promovendo práticas mais eficientes e sustentáveis.

## 2 Referencial Teórico e Metodológico

Este capítulo aborda dois referenciais que formarão a base de nossa discussão, a saber: ontologias e a cultura do tabaco.

Uma ontologia define um conjunto de primitivas representacionais com as quais modelar um domínio de conhecimento ou discurso. As primitivas representacionais são tipicamente classes (ou conjuntos), atributos (ou propriedades) e relacionamentos (ou relações entre membros de classes). As definições das primitivas representacionais incluem informações sobre seu significado e restrições quanto à sua aplicação lógica e consistente (GRUBER, 1995).

A partir da evidente proximidade entre a lógica ontológica e a teoria dos conjuntos, observa-se que as classes podem ser interpretadas como conjuntos, uma vez que ambas representam coleções de elementos que compartilham características comuns. De forma análoga, os indivíduos correspondem aos elementos, sendo membros específicos das classes ou conjuntos que os definem. Ademais, os relacionamentos na ontologia podem ser compreendidos como relações na teoria dos conjuntos, ou seja, conectam indivíduos/elementos intra ou entre classes/conjuntos, e, finalmente, atributos são funções (ou seja, descrevem uma propriedade e dão valor à ela. A Tabela 2.1 ilustra essa conexão.

<b>Ontologia</b>	<b>Teoria dos Conjuntos</b>
Classes	Conjuntos
Indivíduos	Elementos
Relacionamentos	Relações
Atributos	Funções

Tabela 2.1 – Conceitos paralelos entre ontologias e conjuntos

### 2.1 Conceituação de ontologias por Gruber

Nesta sessão, exploramos os princípios fundamentais para o design de ontologias voltadas ao compartilhamento de conhecimento, conforme proposto por Gruber (GRUBER, 1995). Esses princípios estabelecem a base teórica para o desenvolvimento da Ontologia da Produção de Tabaco (OnTop), orientando as decisões de modelagem e estruturação adotadas neste trabalho.

Um corpo de conhecimento formalmente representado é baseado em uma conceituação, ou seja, uma visão de mundo simplificada e abstrata, que modela os conceitos

que desejemos representar. Uma ontologia seria, então, uma especificação explícita dessa conceituação. Isso significa que é uma representação formal dos conceitos, objetos e relações que existem em um determinado domínio de conhecimento. As ontologias permitem que diferentes sistemas e agentes compartilhem entendimento sobre um domínio específico, facilitando a interoperabilidade e a reutilização de conhecimento.

Gruber enfatiza que as ontologias devem ser projetadas com um propósito claro, sendo consideradas artefatos desenvolvidos para atender a objetivos específicos e avaliadas de acordo com critérios objetivos de design. No contexto da produção de tabaco, a necessidade de uma ontologia surge da complexidade inerente ao domínio, que envolve aspectos ambientais, agronômicos, econômicos e sociais. OnTop busca formalizar essa complexidade, proporcionando uma estrutura para representar e compartilhar conhecimento de forma eficaz.

Um dos princípios centrais propostos por Gruber é o da **clareza**. A ontologia deve comunicar efetivamente o significado pretendido dos termos definidos, utilizando definições objetivas e, sempre que possível, formalizadas em lógica. Isso evita ambiguidades e interpretações errôneas dos conceitos. Ao desenvolver OnTop, asseguramos que cada classe e propriedade/atributo fosse claramente definida.

Outro princípio fundamental é o da **coerência**. A ontologia deve ser internamente consistente, permitindo inferências que sejam alinhadas com as definições estabelecidas. Isso implica que os axiomas e restrições lógicas devem ser cuidadosamente elaborados para evitar contradições. Na OnTop, utilizamos ferramentas de raciocínio automatizado para verificar a consistência lógica, garantindo que as relações entre classes, como *precedes* e *follows*, estejam corretamente aplicadas.

A **extensibilidade** é outro princípio destacado por Gruber. A ontologia deve ser projetada de modo a permitir extensões e refinamentos futuros sem a necessidade de reestruturar os componentes existentes. Isso é crucial para acomodar evoluções no domínio de conhecimento ou mudanças nas necessidades dos usuários. Na OnTop, adotamos uma estrutura modular, onde classes genéricas podem ser especializadas conforme necessário.

Gruber também aborda o princípio do **mínimo viés de codificação**. A ontologia deve ser especificada no nível do conhecimento, evitando dependências de representações específicas ou detalhes de implementação. Isso assegura que a ontologia seja independente de tecnologias ou plataformas particulares, facilitando sua adoção em diferentes sistemas. Ao desenvolver a OnTop, evitamos a inclusão de detalhes de implementação específicos, como formatos de dados proprietários ou unidades de medida fixas. Representamos quantidades de forma abstrata, permitindo que diferentes sistemas de unidades sejam utilizados conforme o contexto regional ou as preferências dos usuários.

O último princípio discutido por Gruber é o do **mínimo comprometimento ontológico**. A ontologia deve fazer as menores suposições possíveis sobre o mundo, fornecendo apenas a estrutura necessária para suportar as atividades de compartilhamento de conhecimento pretendidas. Isso permite que diferentes agentes utilizem a ontologia de acordo com suas necessidades, sem serem restringidos por compromissos desnecessários. Na OnTop, incluímos apenas os conceitos essenciais para modelar a produção de tabaco, evitando sobrecarregar a ontologia com detalhes excessivamente específicos. Por exemplo, utilizamos propriedades opcionais para acomodar variações regionais nas práticas agrícolas ou nas classificações de solo.

A aplicação dos princípios de Gruber na OnTop poder resultar em uma ontologia clara, coerente, extensível, livre de vieses de codificação e com mínimo comprometimento ontológico. Isso não apenas facilita o compartilhamento e a reutilização de conhecimento, mas também contribui para a interoperabilidade entre sistemas e agentes no domínio agrícola.

É importante destacar que a formalização de ontologias também desempenha um papel crucial na integração com tecnologias emergentes, como modelos de linguagem extensos (LLMs). A estrutura clara e formalizada da OnTop permite que esses modelos utilizem o conhecimento representado de maneira eficaz, melhorando a precisão e a relevância das respostas em aplicações relacionadas à produção de tabaco.

Além disso, a adoção de ontologias pode promover avanços significativos na agricultura de precisão. Ao representar formalmente informações sobre condições climáticas, tipos de solo, práticas de manejo e outros fatores críticos, é possível desenvolver sistemas de recomendação mais eficazes e personalizados, auxiliando os agricultores na tomada de decisões informadas.

Em suma, os princípios de design de ontologias propostos por Gruber fornecem um arcabouço sólido para o desenvolvimento de ontologias voltadas ao compartilhamento de conhecimento. A OnTop exemplifica a aplicação desses princípios no domínio da produção de tabaco, demonstrando como uma ontologia pode contribuir para a melhoria das práticas agrícolas.

A compreensão e a aplicação desses princípios são fundamentais não apenas para o desenvolvimento da OnTop, mas também para futuras iniciativas que busquem utilizar ontologias como ferramenta para integração e compartilhamento de conhecimento em diferentes domínios. Ao aderir a esses princípios, os desenvolvedores de ontologias podem garantir que seus modelos sejam úteis, adaptáveis e sustentáveis a longo prazo.

## 2.2 Síntese dos Fundamentos da Produção de Tabaco

Considerando que a produção de tabaco para rendimento comercial é uma atividade consideravelmente complexa, nesta sessão busca-se contextualizar de maneira sintética aspectos fundamentais da cultura com base em (EMATER-AL *et al.*, 1981) e em (HOFFMANN; MARCON, 2012).

A cultura do tabaco é uma atividade agrícola que envolve diversas etapas, desde a sementeira até o processo de cura e preparação das folhas para comercialização. A seguir, apresentamos uma breve descrição de cada uma dessas etapas.

A primeira fase do cultivo do tabaco inicia-se com a sementeira. As sementes de tabaco, que são extremamente pequenas, são semeadas em viveiros chamados de canteiros de mudas. Esses canteiros são preparados com solo rico em nutrientes e devidamente protegido de condições climáticas adversas, como geadas e chuvas intensas. As sementes são espalhadas de forma uniforme na superfície do solo e, em seguida, são cobertas levemente com uma fina camada de terra ou material orgânico para garantir a germinação adequada.

Após a sementeira, é necessário garantir que os canteiros recebam uma quantidade adequada de água e luz solar. A irrigação é feita com cuidado para não deslocar as sementes, utilizando sistemas de irrigação fina. A germinação ocorre entre 7 a 10 dias, e as mudas começam a emergir da superfície do solo. Durante esse período, é fundamental monitorar a saúde das mudas e evitar o surgimento de doenças e pragas que possam comprometer o desenvolvimento inicial das plantas.

Com o crescimento das mudas, chega-se à fase de repicagem ou transplante, que ocorre quando as mudas atingem um certo tamanho e resistência adequados para serem transferidas para o campo sem o risco de grandes perdas (pequenas perdas, no entanto, sempre ocorrem). O transplante geralmente ocorre entre 40 e 60 dias após a sementeira. Antes de transferir as mudas, o solo do campo é preparado com práticas de aração e adubação, garantindo uma estrutura adequada para o desenvolvimento das plantas. O espaçamento entre as mudas no campo é cuidadosamente calculado para permitir que cada planta tenha espaço suficiente para crescer e receber nutrientes e luz solar de forma adequada. Maiores detalhes sobre os efeitos negativos de um espaçamento insuficiente entre plantas podem ser encontrados em (LEONE, 1952).

Uma vez transplantadas, as plantas de tabaco entram em fase de crescimento vegetativo, onde é necessário o controle de ervas daninhas, pragas e doenças. Os agricultores utilizam técnicas de capina e rotação de culturas para manter o solo saudável e minimizar o impacto de patógenos. A rotação de culturas é particularmente apropriada para o tabaco, que depleta significativamente do solo tanto nutrientes primários ( $N$ ,  $P$ ,

*K*) quanto secundários (*Mg*, *Ca*). Esse, na realidade, é um problema comum no cultivo de plantas da família *Solanaceae*. A aplicação de defensivos agrícolas também pode ser necessária, e é regida por normas ambientais e de segurança alimentar.

Após algumas semanas de crescimento no campo (em geral de 4 a 6 semanas, a depender das condições e também do cultivar), as plantas de tabaco passam pelo processo de desbrota ou capação. Essa prática consiste em remover a parte superior da planta, onde floresceriam as inflorescências, com o objetivo de forçar a planta a concentrar os nutrientes no crescimento das folhas. A desbrota promove o aumento da qualidade das folhas, uma vez que a planta direciona seus recursos para o desenvolvimento foliar em vez de investir na produção de sementes.

Após a desbrota, ocorre o desenvolvimento completo das folhas, que, em seguida, estão prontas para a colheita. O processo de colheita do tabaco é feito em etapas, colhendo-se as folhas conforme atingem o ponto ideal de maturação. No caso do cultivar Virgínia, por exemplo, as folhas são colhidas em até quatro fases, começando pelas folhas inferiores (mais próximas ao solo), que amadurecem primeiro, até chegar às folhas superiores. Cada fase da colheita tem características próprias, já que as folhas apresentam diferentes qualidades e são destinadas a diferentes tipos de processamento.

Depois de colhidas, as folhas passam pelo processo de cura, que pode ocorrer em estufas ou galpões especialmente construídos para esse fim. A cura é um processo crucial para o desenvolvimento do sabor, cor e aroma das folhas de tabaco. Existem diferentes tipos de cura, sendo as mais comuns a cura ao ar e a cura ao calor. A cura ao ar (utilizada para os cultivares Burley e Maryland, por exemplo), envolve a secagem lenta das folhas em galpões abertos, onde as folhas são penduradas e secam naturalmente com a circulação do ar. Já a cura ao calor, utilizada para os cultivares Virgínia e Amarelinho, envolve o uso de estufas que controlam a temperatura e a umidade, permitindo uma secagem mais rápida e uniforme das folhas.

Durante o processo de cura, é necessário monitorar constantemente a temperatura e a umidade no interior dos galpões ou estufas, uma vez que variações excessivas podem comprometer a qualidade final do produto. A mofagem é um problema sério nesse estágio da produção. A duração da cura pode variar de acordo com o cultivar e as condições climáticas, mas geralmente dura entre uma a três semanas.

Após a cura, as folhas de tabaco passam pelo processo de classificação, onde são separadas por qualidade, tamanho e coloração. A classificação é uma etapa importante, pois define o valor comercial das folhas. As folhas são cuidadosamente analisadas e agrupadas de acordo com suas características físicas, sendo, então, preparadas para o armazenamento ou para o processo de fermentação.

Em alguns tipos de tabaco, como o tabaco para charutos, as folhas passam por um processo adicional de fermentação, que realça ainda mais o sabor e o aroma. A fermentação é um processo controlado de decomposição enzimática das folhas, que ocorre em ambientes específicos com controle de temperatura e umidade. Esse processo pode durar vários meses e é fundamental para a produção de tabacos de alta qualidade.

Finalmente, após todos os processos de cura, fermentação e classificação, as folhas de tabaco são embaladas e preparadas para a comercialização. O tabaco é armazenado em condições controladas para preservar sua qualidade até o momento de ser utilizado.

## 3 Revisão de Literatura e Trabalhos Relacionados

Neste capítulo, apresentamos uma síntese da revisão sistemática da literatura (ALVES; FRANCO, 2024) que foi conduzida com foco em sistemas de recomendação baseados em Sistemas de Informação Geográfica (GIS) aplicados à agricultura. Essa pesquisa representa o estágio inicial de nossa investigação, os resultados da qual apontaram para a possibilidade e o interesse em uma ontologia em agricultura. A revisão fornece uma visão abrangente sobre o estado da arte nessa área. Além disso, discutimos trabalhos relacionados, incluindo ontologias existentes no domínio do tabaco, destacando as contribuições e limitações de cada abordagem.

A agricultura moderna enfrenta desafios crescentes devido à necessidade de aumentar a produção alimentícia de forma sustentável e eficiente (ED-DAOUDI *et al.*, 2023). Nesse contexto, sistemas de recomendação baseados em GIS surgem como ferramentas promissoras para auxiliar agricultores na tomada de decisões informadas, considerando variáveis espaciais e ambientais (PIERCE *et al.*, 2007). Tais sistemas combinam dados geoespaciais com técnicas de inteligência artificial para oferecer recomendações personalizadas sobre práticas agrícolas, seleção de culturas, manejo de solo, entre outros aspectos.

### 3.1 Sistemas de Recomendação Baseados em GIS na Agricultura

Alves & Rosa (2024) (ALVES; FRANCO, 2024) apresentam uma revisão sistemática da literatura sobre sistemas de recomendação baseados em GIS aplicados à agricultura. O estudo teve como objetivo analisar as tendências atuais, identificar lacunas de pesquisa e propor direções futuras para o desenvolvimento de sistemas eficazes nesse domínio.

#### 3.1.1 Metodologia da Revisão Sistemática

Na revisão sistemática, foram empregadas as diretrizes de Kitchenham (KITCHENHAM, 2004). Critérios de inclusão e exclusão foram definidos e aplicados. Foram consideradas publicações entre 2018 e 2023, abrangendo artigos de revistas e de conferências. A busca foi realizada em bases de dados reconhecidas (IEEE Xplore e Springer Link).

Considerando palavras-chave específicas relacionadas a sistemas de recomendação, GIS e agricultura, a seguinte string de busca foi utilizada:

*(gis OR geographical OR spatial) AND (agriculture OR farming OR farm OR agribusiness OR crop OR cultivation OR culture OR harvest OR fertilizer OR soil) AND (recommender OR recommendation OR suggestion)*

Após a aplicação dos critérios, 26 artigos foram selecionados para a análise detalhada. Os trabalhos foram categorizados de acordo com os tipos de culturas abordadas, técnicas utilizadas e os principais objetivos dos sistemas propostos.

### 3.1.2 Resultados e Discussão

A revisão revelou que a maioria dos sistemas de recomendação baseados em GIS na agricultura concentra-se em culturas alimentícias, como arroz, trigo e milho. Poucos estudos abordam culturas específicas, não-alimentícias ou de alto valor econômico, como o tabaco. Além disso, foi observado que técnicas de aprendizado de máquina, como árvores de decisão (do inglês *Decision Trees - DT*), máquinas de vetor de suporte (do inglês *Support Vector Machines - SVM*) e florestas aleatórias (do inglês *Random Forests - RF*), são comumente utilizadas para analisar dados geoespaciais e gerar recomendações.

Os principais objetivos dos sistemas analisados incluem a otimização da seleção de culturas (FARHEEN, 2021), (ANLEY; TESEMA, 2019), (DOSHI *et al.*, 2018), (SUCHITHRA; PAI, 2020), (PRIYADHARSHINI *et al.*, 2021), (RAMCHANDANI *et al.*, 2020), (ED-DAOUDI *et al.*, 2023), (EZRA *et al.*, 2023), a avaliação da fertilidade do solo (RAJAMANI *et al.*, 2020), (SUCHITHRA; PAI, 2020), (RAMCHANDANI *et al.*, 2020), (RAHMAN *et al.*, 2018), (PRASAD *et al.*, 2021), (AVEZBOYEV *et al.*, 2023), (YUDHANA *et al.*, 2023), (NUNGULA *et al.*, 2023) e a melhoria do manejo de pragas e doenças (RODRÍGUEZ-GARCÍA; GARCÍA-SÁNCHEZ, 2020), (GAO *et al.*, 2018).

Muitos trabalhos destacaram a importância da integração de dados ambientais e climáticos para aumentar a precisão das recomendações (ANLEY; TESEMA, 2019), (FARHEEN, 2021), (GAO *et al.*, 2018), (TINOCO *et al.*, 2019), (ZERMAS *et al.*, 2021), (RIZALDI *et al.*, 2019), (OSMAN *et al.*, 2020), (PETERS *et al.*, 2020).

A Tabela 3.1 apresenta uma síntese dos estudos analisados, incluindo autor, ano, objetivo, técnicas utilizadas e apresentação.

Tabela 3.1 – Síntese comparativa dos estudos analisados

Referências	T-THE	T-APP	G-MOD	G-MNG	U-SSL	U-CRE	U-FER	U-PCT	U-WAT	U-YPD	GIS-C	GIS-S	GIS-O	P-DES	P-MAT
(ANLEY; TESEMA, 2019)	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0
(CHEBOTAREV <i>et al.</i> , 2020)	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0
(DOSHI <i>et al.</i> , 2018)	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0
(FARHEEN, 2021)	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0
(GAO <i>et al.</i> , 2018)	0	1	1	0	0	0	1	1	1	0	1	0	0	1	0
(IYER <i>et al.</i> , 2019)	0	1	0	1	1	0	1	0	0	0	0	0	1	0	1
(MUPANGWA <i>et al.</i> , 2020)	0	1	0	1	0	0	0	0	0	1	0	0	1	0	1
(SUCHITHRA; PAI, 2020)	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0
(TINOCO <i>et al.</i> , 2019)	0	1	0	1	1	0	0	0	0	0	1	0	0	0	1
(YUSIANTO <i>et al.</i> , 2020)	0	1	1	0	0	0	1	0	0	0	0	0	1	0	1
(ZERMAS <i>et al.</i> , 2021)	0	1	0	1	0	0	1	0	0	0	1	0	0	0	1
(DABRE <i>et al.</i> , 2018)	1	0	1	0	0	0	1	0	1	0	0	0	1	1	0
(FEGADE; PAWAR, 2020)	1	0	1	0	0	0	0	0	0	1	0	0	1	1	0
(PRIYADHARSHINI <i>et al.</i> , 2021)	1	0	1	0	0	1	0	0	0	0	1	0	0	0	1
(RAMCHANDANI <i>et al.</i> , 2020)	1	0	1	0	0	1	0	0	0	0	0	1	0	0	1
(RIZALDI <i>et al.</i> , 2019)	1	0	1	0	1	0	1	0	0	0	1	0	0	0	1
(RAHMAN <i>et al.</i> , 2018)	1	0	0	1	1	0	1	0	0	0	0	1	0	0	1
(OSMAN <i>et al.</i> , 2020)	1	0	0	1	0	0	0	0	0	1	1	0	0	1	0
(PETERS <i>et al.</i> , 2020)	1	0	0	1	0	0	0	0	0	1	1	0	0	1	0
(RODRÍGUEZ-GARCÍA; GARCÍA-SÁNCHEZ, 2020)	1	0	0	1	0	0	0	1	0	0	1	0	0	1	0
(PRASAD <i>et al.</i> , 2021)	0	1	0	1	0	0	1	0	0	0	0	1	0	1	0
(ED-DAOUDI <i>et al.</i> , 2023)	0	1	1	0	0	1	0	0	0	0	0	1	0	1	0
(AVEZBOYEV <i>et al.</i> , 2023)	0	1	0	1	1	0	0	0	0	0	0	1	0	1	0
(EZRA <i>et al.</i> , 2023)	1	0	0	1	1	1	0	0	0	0	1	0	0	1	0
(YUDHANA <i>et al.</i> , 2023)	0	1	0	1	1	0	0	0	0	0	0	1	0	0	1
(NUNGULA <i>et al.</i> , 2023)	0	1	1	0	1	0	0	0	0	0	0	1	0	1	0

As categorias utilizadas na Tabela 3.1 são as seguintes:

### Tipo (Type)

- **T-THE (Teórico)**: se o artigo apresenta uma estrutura teórica e descreve uma possível aplicação.
- **T-APP (Aplicação)**: se o artigo apresenta uma estrutura teórica e a aplica em um cenário do mundo real.

### Objetivo (Goal)

- **G-MOD (Modernização)**: se o objetivo do artigo é fornecer meios de modernização para práticas agrícolas que não correspondem ao estado moderno do negócio.
- **G-MNG (Gestão)**: se o artigo apresenta uma ferramenta para melhorar processos que já estão em um nível maduro de mecanização ou informatização.

### Uso no Ciclo de Vida da Cultura (Use on the Crop Lifecycle)

- **U-SSL (Seleção de Solo)**: se o artigo trata das características do solo, sem considerar outros elementos do processo agrícola.
- **U-CRE (Recomendação de Cultura)**: se a abordagem proposta recomenda uma cultura, dadas as possibilidades da terra arável.
- **U-FER (Fertilização)**: se a abordagem sugere fertilizantes (por exemplo, NPK) para lidar com deficiências específicas do campo.
- **U-PCT (Controle de Pragas)**: se a abordagem recomenda pesticidas adequados para uma cultura específica e para a qualidade do solo.
- **U-WAT (Irrigação)**: se a abordagem sugere o volume e o cronograma de irrigação para a agricultura de produção.
- **U-YPD (Previsão de Rendimento)**: se a abordagem estima o rendimento da cultura, considerando os elementos das fases anteriores do ciclo de vida.

### Uso de GIS (GIS use)

- **GIS-C (Clima)**: se na abordagem o uso de GIS foi principalmente voltado para lidar com precipitação, temperatura, eventos climáticos e outros assuntos específicos do clima.

- **GIS-S (Saúde do Solo)**: se na proposta o GIS foi usado para mapear as características do solo, como o tipo de solo ou a concentração de macronutrientes.
- **GIS-O (Compreensivo)**: o uso de GIS abrange tanto **Clima** quanto **Saúde do Solo**.

### Apresentação (Presentation)

- **P-DES (Descritiva)**: se o trabalho foi descrito de forma qualitativa.
- **P-MAT (Matemática)**: se os aspectos matemáticos do método estão claramente expostos no artigo, permitindo diferentes níveis de reprodução.

A partir da revisão, identificou-se que, apesar dos avanços, ainda existem lacunas significativas na aplicação de sistemas de recomendação baseados em GIS para culturas específicas. A maioria dos estudos foca em culturas alimentares, deixando em segundo plano culturas de grande importância econômica em determinadas regiões.

Nota-se também a falta de integração entre ontologias e sistemas de recomendação, o que poderia melhorar a representação do conhecimento e a interoperabilidade entre sistemas (GRUBER, 1995). A utilização de ontologias de domínio específicas permite uma modelagem mais precisa dos conceitos e relações inerentes à agricultura, facilitando a troca de informações e o desenvolvimento de soluções mais eficazes. De fato, nenhum dos artigos estudados apresenta uma ontologia.

## 3.2 Trabalhos Relacionados

Nesta seção, discutimos 2 trabalhos considerados relacionados por apresentarem ontologias no domínio do tabaco. Uma síntese de cada trabalho é apresentada, além de uma análise comparativa com a proposta apresentada nesta dissertação. A seleção desses artigos foi feita buscando-se títulos que contivessem ambos "tobacco" e "ontology".

(COX *et al.*, 2023) propuseram uma ontologia voltada para produtos de tabaco, especificamente nicotina e cigarros eletrônicos. O objetivo foi criar uma estrutura que permitisse a padronização de termos e conceitos relacionados a produtos de tabaco e dispositivos de vaping, facilitando pesquisas e políticas públicas nessa área.

A ontologia desenvolvida por (COX *et al.*, 2023) concentra-se nos aspectos de consumo e nos produtos em si, detalhando características como componentes químicos, tipos de dispositivos e padrões de uso. Embora seja uma contribuição relevante, essa

ontologia não aborda aspectos relacionados à produção agrícola do tabaco, como práticas de cultivo, manejo de solo ou condições ambientais.

(NOTLEY *et al.*, 2022) desenvolveram uma ontologia focada em construtos relacionados à identidade no contexto de vícios, com ênfase no uso de nicotina e tabaco. A ontologia visa esclarecer como a identidade individual e social influencia comportamentos de vício, auxiliando em pesquisas sobre dependência e estratégias de cessação.

Semelhante ao trabalho de Cox *et al.*, a ontologia de Notley *et al.* concentra-se nos aspectos psicológicos e comportamentais associados ao consumo de tabaco. Não há, entretanto, uma representação das etapas de produção ou dos fatores agrônômicos envolvidos na cultura do tabaco.

A análise dos trabalhos de (COX *et al.*, 2023) e (NOTLEY *et al.*, 2022) evidencia uma lacuna na representação formal do conhecimento relacionado à produção de tabaco. Enquanto essas ontologias contribuem de maneira efetiva para os estudos de consumo, vício e regulamentação de produtos de tabaco, não abordam os processos agrícolas e ambientais que antecedem o consumo.

Por outro lado, ao analisar as ontologias agrícolas disponíveis no Crop Ontology Curation Tool <sup>1</sup>, observamos que elas abrangem uma variedade de culturas alimentares e forrageiras, como arroz, milho, trigo e soja. No entanto, não existem ali ontologias específicas para o tabaco, uma cultura de grande relevância econômica. Além disso, as ontologias existentes concentram-se principalmente em aspectos agrônômicos e fenotípicos, sem abordar as particularidades do cultivo do tabaco, como práticas de manejo específicas, processos de cura das folhas e fatores ambientais que influenciam a qualidade final do produto e que são demasiado particulares dessa cultura. Diante dessa lacuna, optamos por focar em ontologias relacionadas ao tabaco, mesmo que não diretamente ligadas à agricultura, para compreender melhor os conceitos e terminologias associadas a essa cultura.

OnTop destina-se a preencher essa lacuna, oferecendo uma ontologia específica para a produção de tabaco que engloba todo o ciclo de vida da cultura, desde o preparo do solo até a cura das folhas. Ao incorporar fatores como condições climáticas, características geográficas, tipos de solo, manejo de pragas e práticas de manejo, OnTop fornece uma estrutura conceitual abrangente para representar e compartilhar conhecimento no domínio agrícola do tabaco.

A revisão da literatura demonstra a crescente importância de sistemas de recomendação na agricultura e destaca a necessidade de ontologias específicas de domínio para aprimorar esses sistemas. Embora existam ontologias relacionadas ao tabaco, nenhuma

---

<sup>1</sup> Crop Ontology Curation Tool - [www.cropontology.org](http://www.cropontology.org)

aborda a produção agrícola em si.

OnTop surge, portanto, como uma contribuição inédita, preenchendo uma necessidade identificada tanto na literatura quanto na prática agrícola. Ao alinhar-se com os princípios de design de ontologias propostos por Gruber (1995) (GRUBER, 1995), OnTop oferece um framework robusto para o desenvolvimento de soluções tecnológicas que auxiliem os produtores de tabaco na tomada de decisões informadas e sustentáveis.

## 4 Ontologia da Produção de Tabaco

A produção de tabaco é um processo complexo que envolve uma série de etapas interdependentes e sequenciais, desde a preparação do solo até a colheita e cura das folhas. A necessidade de formalizar e representar esse conhecimento de maneira estruturada motivou o desenvolvimento da Ontologia da Produção de Tabaco (*Ontology of Tobacco Production - OnTop*) (ALVES *et al.*, 2024). Neste capítulo, exploramos em detalhes os aspectos teóricos que fundamentam a OnTop. Para facilitar a identificação, classes e subclasses são destacadas em **negrito** e *itálico*.

OnTop foi concebida para preencher uma lacuna na representação formal do conhecimento relacionado à produção de tabaco. Apesar da importância econômica e social dessa cultura em diversos países, não existia até então uma ontologia que modelasse todo o ciclo da produção de tabaco. A criação de OnTop tem como objetivo oferecer uma estrutura que permita a integração de informações, a interoperabilidade entre sistemas e o suporte à tomada de decisão baseada em conhecimento consolidado.

O desenvolvimento de OnTop baseou-se nos princípios de design de ontologias propostos por Gruber (GRUBER, 1995), garantindo clareza, coerência, extensibilidade, mínimo viés de codificação e mínimo comprometimento ontológico. A ontologia foi implementada utilizando a Web Ontology Language (OWL), uma linguagem padrão para a criação de ontologias na Web, que permite a representação de conceitos, propriedades e relações de maneira formal e compatível com ferramentas de raciocínio automatizado. A Versão 1.0 de OnTop (OnTop\_V1.0.owl) está disponível no repositório GitHub (ALVES; ROSA, 2024).

O editor de ontologias utilizado foi o *Protégé*, que é o editor de ontologias mais popular do mundo. *Protégé* foi lançado há mais de 25 anos e foi desenvolvido pelo *Stanford Center for Biomedical Informatics Research* (MUSEN, 2015). Ele fornece uma interface gráfica de fácil utilização para a criação de ontologias, além de disponibilizar diversas ferramentas em seu ecossistema, como *reasoners* e ferramentas de criação e visualização de estrutura de grafos. A Figura 4.1 mostra a tela inicial do referido logiciário após o carregamento do arquivo contendo a OnTop.

OnTop modela o ciclo da produção agrícola do tabaco por meio de classes que representam conceitos fundamentais do domínio. As principais classes (***ClimateCondition***, ***HumanAgent***, ***ManagementAction***, ***PestsAndDiseases***, ***ProductionStage***, ***SoilCondition***, ***TobaccoField***, ***TobaccoLifecycle***, ***YieldOutcome***) definem os conceitos centrais para a produção do tabaco: elas são os pilares sobre os quais será

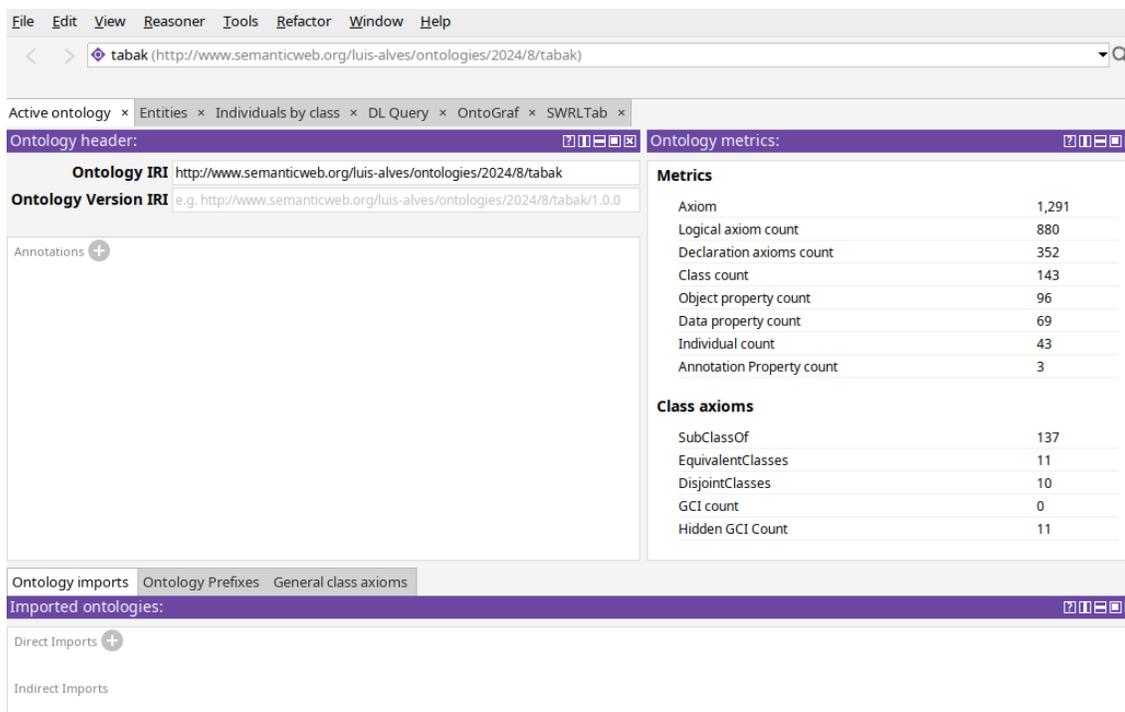


Figura 4.1 – Tela inicial do logiciário *Protégé*

construída a ontologia. Um erro na escolha dessas classes principais pode vir a criar grandes problemas lógicos no futuro. Cada classe é definida com base em propriedades e restrições que capturam as características essenciais dos conceitos representados.

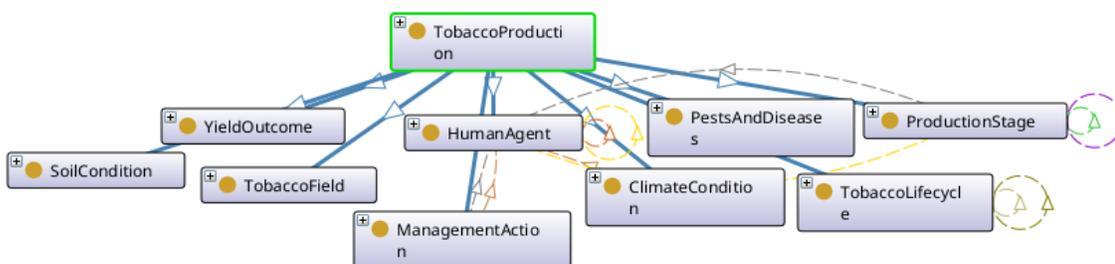


Figura 4.2 – Classes Principais de OnTop.owl

A estrutura central do OnTop é composta por 9 classes principais, cada uma representando um aspecto fundamental da produção de tabaco. Essas classes, que herdam diretamente de ***TobaccoProduction***, servem como a estrutura base para modelar componentes e processos no domínio da agricultura de tabaco. A seguir são apresentadas as definições das classes principais de OnTop, conforme proposto em (ALVES *et al.*, 2024)

A classe ***ClimateCondition*** engloba fatores ambientais críticos como ***Temperature***, ***Humidity***, ***RainfallAmount***, ***SunlightExposure***, ***FrostIncidence*** e ***HailstoneOccurrence***. Esses elementos influenciam o crescimento e a saúde das plantas de tabaco, afetando decisões sobre épocas de plantio, irrigação e controle de pragas. Cada

subclasse fornece informações detalhadas sobre atributos climáticos específicos, permitindo uma modelagem precisa dos impactos ambientais no cultivo de tabaco.

**ManagementAction** representa várias atividades e intervenções realizadas durante a produção de tabaco. Esta classe inclui atividades como **FertilizerApplication**, **PesticideApplication**, **Irrigation**, **Weeding**, **SoilPreparation** e **LimeApplication**. Ao formalizar essas etapas, OnTop apoia a otimização das práticas agrícolas, permitindo um melhor planejamento e alocação de recursos. A representação estruturada dessas atividades de manejo também ajuda a avaliar sua eficácia no rendimento e na qualidade das colheitas.

**PestsAndDiseases** aborda fatores de estresse biótico que podem afetar negativamente a produção de tabaco. Inclui subclasses como **IncidenceOfPestsAndDiseases** e **NoxiousWeeds**, detalhando ameaças específicas à colheita de tabaco. Essa classificação é crucial para integrar estratégias de manejo de pragas e doenças no quadro mais amplo de suporte à decisão, garantindo que as medidas de mitigação sejam oportunas e eficazes.

**ProductionStage** detalha as 5 fases do ciclo de vida do tabaco, i.e., **TobaccoSeedbed**, **TobaccoPlanting**, **TobaccoTopping**, **TobaccoHarvest** e **TobaccoCuring**. Essas etapas são sequenciais e interdependentes, levando em consideração requisitos específicos de manejo e considerações ambientais. Essa classe e suas subclasses visam a facilitar o planejamento detalhado e o monitoramento ao longo do processo de produção, garantindo que cada etapa seja executada de forma otimizada para maximizar o rendimento e a qualidade.

**TobaccoLifecycle** oferece uma visão abrangente do crescimento e desenvolvimento das plantas de tabaco, englobando as entidades **TobaccoSeed**, **TobaccoSeedling**, **TobaccoPlant** e **TobaccoLeaf**. Essa classe é essencial para modelar aspectos biológicos e fenológicos da produção de tabaco, permitindo uma compreensão holística da dinâmica de crescimento da cultura. Ela interage de perto com outras classes principais, como **SoilCondition**, ilustrando a complexa interação entre fatores biológicos e ambientais no cultivo do tabaco.

**HumanAgent** representa os indivíduos envolvidos na produção de tabaco, como agricultores, consultores agrícolas e agentes de seguros. Esta classe é crucial para capturar os papéis e responsabilidades dos atores na gestão e suporte ao ciclo de produção de tabaco. Ao formalizar as interações e decisões dos agentes humanos, OnTop permite uma melhor compreensão de como os fatores humanos influenciam os resultados agrícolas e a implementação das melhores práticas no cultivo de tabaco.

**SoilCondition** modela as propriedades do solo que são críticas para o cultivo

de tabaco, incluindo **pH**, **SoilMoisture** e **SoilType**. Essas características são essenciais para determinar a adequação do solo para a produção de tabaco e para tomar decisões informadas sobre práticas de manejo do solo (e.g., fertilização e irrigação). A integração das condições do solo em nossa ontologia permite recomendações mais precisas para otimizar a saúde das plantas e o rendimento.

**TobaccoField** abrange os aspectos espaciais e geográficos do cultivo de tabaco, incluindo a **GeographicalLocation**, **Size** e **Soil** de cada campo. Esta classe é fundamental para vincular os atributos físicos dos campos de tabaco com dados ambientais e de manejo, possibilitando recomendações e intervenções específicas para cada local. Ao representar as características únicas de cada campo, nossa ontologia apoia a aplicação de técnicas de agricultura de precisão adaptadas a parcelas individuais.

Por fim, a Classe **YieldOutcome** quantifica os resultados da produção de tabaco, representando métricas como rendimento por hectare (**YieldPerHectare**) e o rendimento total (**YieldOutcome**). Esta classe é essencial para avaliar a eficácia de várias práticas de manejo e condições ambientais sobre o rendimento de tabaco. Ao integrar os dados de produção com outros elementos ontológicos, nossa ontologia fornece uma estrutura robusta para avaliar o impacto de diferentes variáveis nos resultados produtivos e orientar a tomada de decisões futuras.

As principais classes e subclasses de OnTop são apresentadas na Figura 4.3, que é uma captura de tela da seção de "Classes" na aba de "Entities" do Protégé. Na opção *Usage* pode-se ver os usos da classe: de quais classes ela é domínio e imagem, com quais classes ela é disjunta, quais são seus indivíduos, quais são suas subclasses, quais as suas propriedades de dados e de objeto, i.e., uma descrição geral do uso dessa classe. O exemplo de uma parte dessa tela é mostrada na Figura 4.4, na classe **TobaccoField**, que modela o talhão onde é plantado o tabaco.

O Protégé permite o uso de *Semantic Web Rule Language* (SWRL) para a criação de regras que serão posteriormente usadas para inferir conhecimento novo a partir dos dados existentes em OWL. SWRL representa uma extensão das capacidades de raciocínio de OWL adicionando regras que vão além da sua expressividade. Regras em SWRL são escritas na forma "*antecedente m implica consequente n*", onde **m** é a condição que precisa ser satisfeita e **n** é a inferência feita no caso da satisfação de **m**. Por exemplo, a condição "*plantas que tem uma altura maior do que 30 unidades (e.g., centímetros) são consideradas plantas adultas*" seria descrita em sintaxe SWRL da seguinte forma:

$$\text{TobaccoPlant}(?p) \wedge \text{hasHeight}(?p, ?\text{height}) \wedge \text{swrlb:greaterThan}(?\text{height}, 30) \rightarrow \text{AdultPlant}(?p)$$

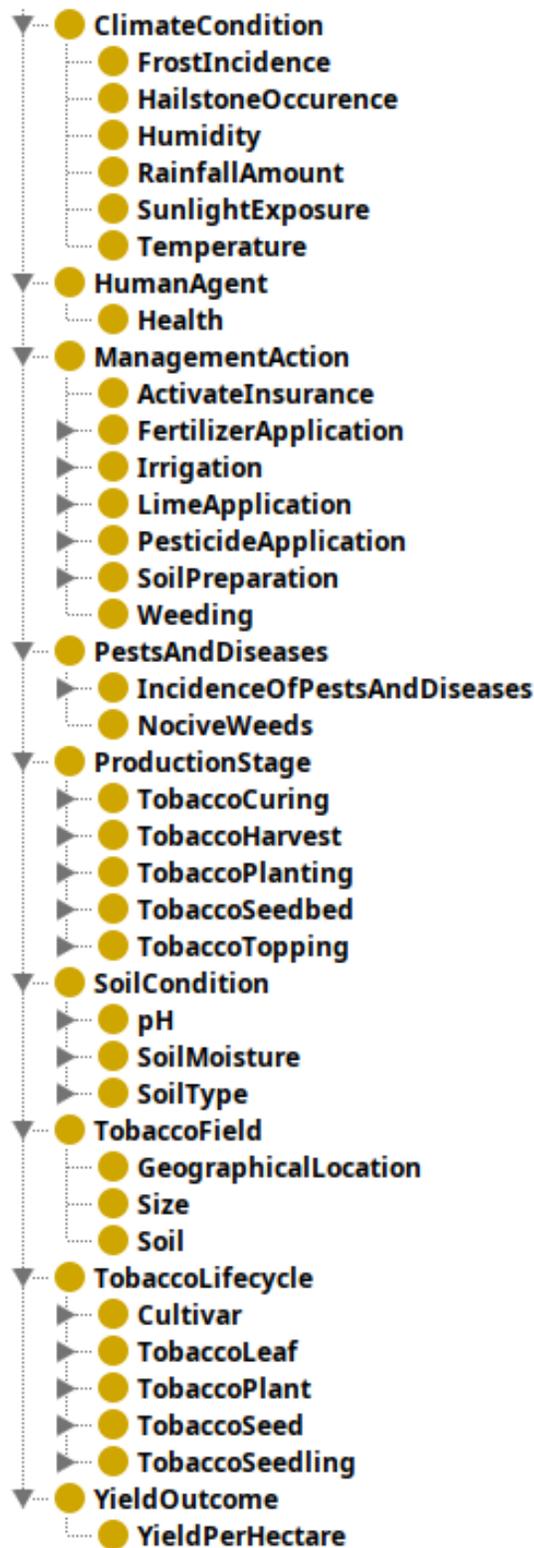


Figura 4.3 – Principais classes e subclasses de OnTop

SWRL permite a definição de regras complexas, a criação de regras para classificação ou tomada de decisão, e ainda permite tirar mais proveito de *reasoners*. *Reasoners*

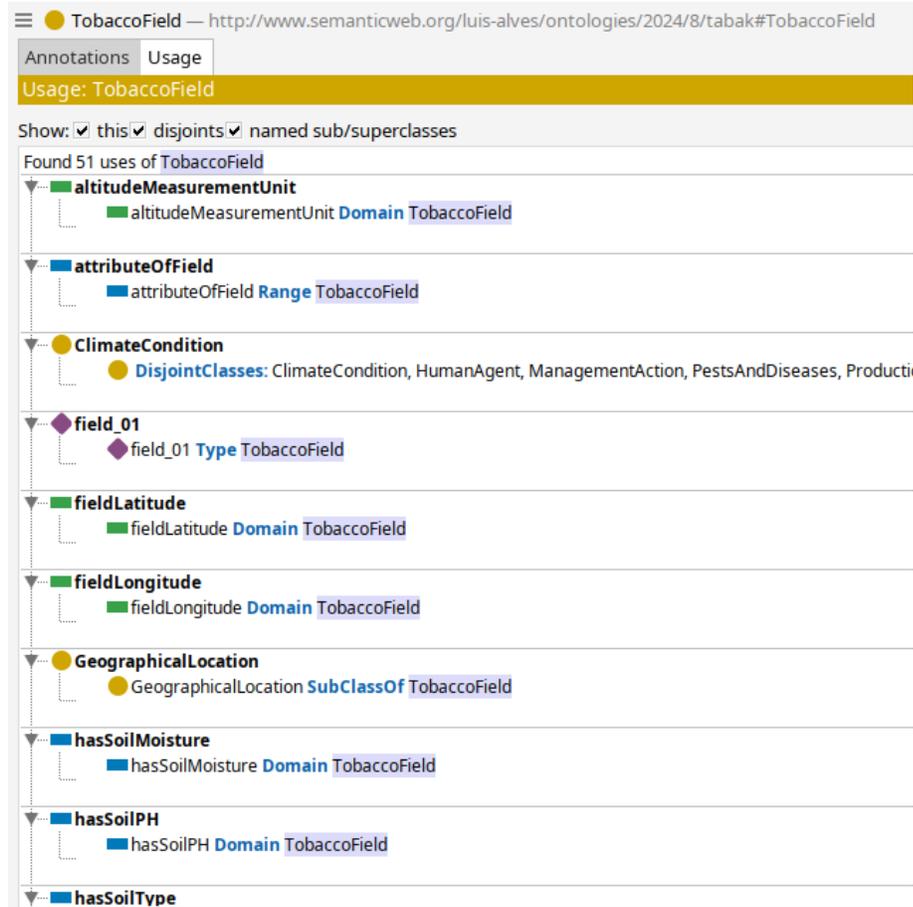


Figura 4.4 – A descrição dos usos de uma classe no *Protégé*

(e.g., HermiT e Pellet) são ferramentas que realizam inferências lógicas com base nas regras e axiomas definidos na ontologia. Eles ajudam a garantir a consistência da ontologia, classificá-la e inferir novos conhecimentos que não estão explicitamente declarados. Essencialmente, eles ampliam as capacidades de raciocínio da ontologia ao analisar as relações entre classes, propriedades e indivíduos para derivar fatos implícitos e identificar inconsistências. Um exemplo do uso de tais regras é mostrado na Figura 4.5.

Active ontology × Entities × Individuals by class × DL Query × OntoGraf × SWRLTab ×		
	Name	
<input checked="" type="checkbox"/>	S1	tabak:CuredLeaf(?x) ^ tabak:isUndercured(?x, true) -> tabak:isOvercured(?x, false)
<input checked="" type="checkbox"/>	S2	tabak:CuredLeaf(?x) ^ tabak:isOvercured(?x, true) -> tabak:isUndercured(?x, false)

Figura 4.5 – Exemplo do uso de SWRL no *Protégé* no contexto de OnTop

## 4.1 Exemplos de Definições Formais

***TobaccoSeedbed*** é uma classe primitiva que representa a etapa de produção relacionada ao viveiro de mudas de tabaco. Esta classe está associada a um tipo específico de tabaco (e.g., Virginia, Connecticut, Burley) por meio da propriedade *hasTobaccoType*, e possui propriedades adicionais que detalham aspectos operacionais, como o número de bandejas plásticas utilizadas para a germinação das sementes, representado pela propriedade *numberOfPlasticTrays*.

Em Lógica Descritiva (DL), podemos definir a classe ***TobaccoSeedbed*** e suas restrições da seguinte forma:

$$\begin{aligned} \text{TobaccoSeedbed} \sqsubseteq & \forall \text{hasClimateCondition. ClimateCondition} \\ & \sqcap \forall \text{hasSoilCondition. SoilCondition} \\ & \sqcap \exists \text{precedes. TobaccoPlanting} \\ & \sqcap \exists \text{hasTobaccoType. TobaccoType} \\ & \sqcap \forall \text{numberOfPlasticTrays. Integer} \end{aligned}$$

Essa definição indica que ***TobaccoSeedbed*** está associada a condições climáticas (*hasClimateCondition*) e de solo (*hasSoilCondition*), possui uma relação de precedência (*precedes*) com ***TobaccoPlanting***, está associada a um tipo de tabaco específico (*hasTobaccoType*) e possui uma propriedade *numberOfPlasticTrays* (que é um número inteiro) representando o número de bandejas plásticas utilizadas.

Utilizando a sintaxe Manchester, a definição acima pode ser expressa como:

Class: TobaccoSeedbed

Annotations:

`rdfs:comment "Etapa de viveiro de mudas na produção de tabaco"`

SubClassOf:

`hasClimateCondition only ClimateCondition,`  
`hasSoilCondition only SoilCondition,`  
`precedes some TobaccoPlanting,`  
`hasTobaccoType some TobaccoType,`  
`numberOfPlasticTrays exactly 1 xsd:integer`

Na definição em Manchester, especificamos que:

- *hasClimateCondition only ClimateCondition*: toda condição climática associada é do tipo *ClimateCondition*.

- *hasSoilCondition only SoilCondition*: toda condição de solo associada é do tipo *SoilCondition*.

- *precedes some TobaccoPlanting*: precede pelo menos uma instância de **TobaccoPlanting**.

- *hasTobaccoType some TobaccoType*: está associado a pelo menos um tipo de tabaco.

- *numberOfPlasticTrays exactly 1 xsd:integer*: possui exatamente um valor inteiro para *numberOfPlasticTrays*.

A propriedade *hasTobaccoType* relaciona **TobaccoSeedbed** a um tipo específico de tabaco. A classe **TobaccoType** pode ser definida como uma enumeração dos diferentes tipos de tabaco cultivados:

Class: TobaccoType

Annotations:

`rdfs:comment "Tipos de tabaco cultivados"`

EquivalentTo:

`{Virginia, Burley, Connecticut}`

Neste exemplo, *Virginia*, *Burley* e *Connecticut* são indivíduos (instâncias) da Classe **TobaccoType**.

A propriedade *numberOfPlasticTrays* é uma propriedade de dados (*data property*) que associa um **TobaccoSeedbed** a um valor inteiro representando o número de bandejas plásticas utilizadas:

DataProperty: numberOfPlasticTrays

Domain: TobaccoSeedbed

Range: xsd:integer

Essa propriedade permite quantificar recursos utilizados na etapa de viveiro, o que é relevante para o planejamento e gerenciamento da produção.

As subclasses **TobaccoSeedbed** e **TobaccoSeed** pertencem a classes diferentes, mas estão intimamente relacionadas de diferentes formas. Essa intrincada relação entre o canteiro onde é plantada a semente e a própria semente pode ser melhor visualizada em suas relações no *Protégé*, através da ferramenta *OntoGraf* (uma ferramenta interna do *Protégé* para a visualização de ontologias), como mostrado na figura 4.6.

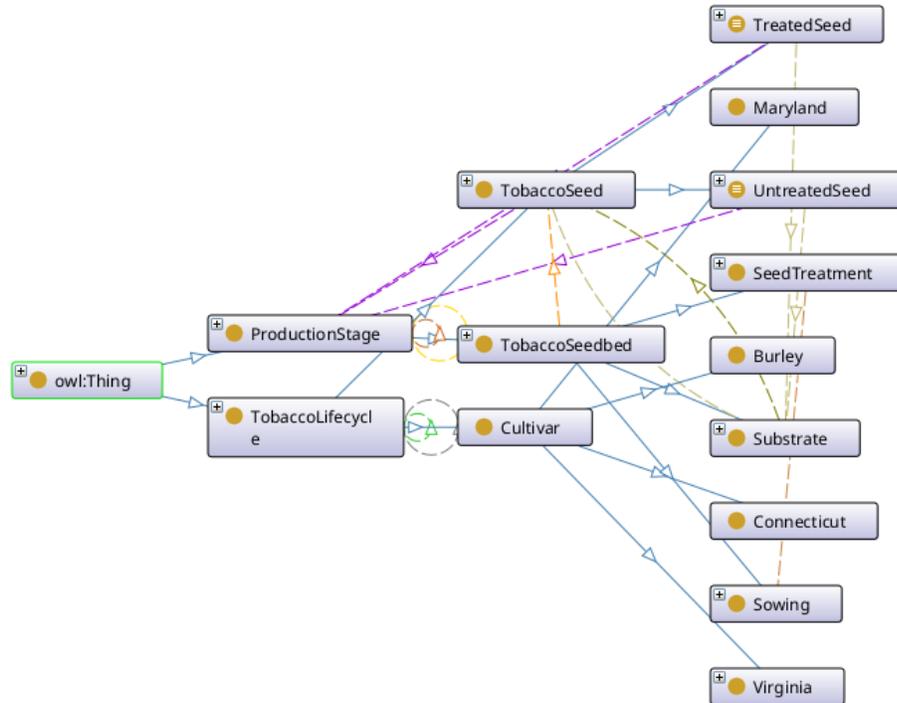


Figura 4.6 – Relação entre *TobaccoSeedbed* e *TobaccoSeed*

De forma similar, podemos definir a Classe *TobaccoPlanting* e suas relações:

$$\begin{aligned} \text{TobaccoPlanting} &\sqsubseteq \forall \text{hasClimateCondition. ClimateCondition} \\ &\sqcap \forall \text{hasSoilCondition. SoilCondition} \\ &\sqcap \exists \text{follows. TobaccoSeedbed} \end{aligned}$$

Em sintaxe Manchester:

```
Class: TobaccoPlanting
Annotations:
  rdfs:comment "Etapa de transplante das mudas de tabaco para o campo"
SubClassOf:
  hasClimateCondition only ClimateCondition,
  hasSoilCondition only SoilCondition,
  follows some TobaccoSeedbed
```

Nessa definição, *TobaccoPlanting* segue (*follows*) pelo menos uma instância de *TobaccoSeedbed*, mantendo a sequência lógica das etapas do processo produtivo.

Ao formalizar OnTop utilizando Lógica Descritiva e a sintaxe Manchester, proporcionamos uma representação precisa e formal do conhecimento, que pode ser utilizada

por *reasoners* para verificar a consistência, inferir novos conhecimentos e auxiliar na tomada de decisão. Essa abordagem fortalece a aplicabilidade da ontologia em sistemas de inteligência artificial e ferramentas de suporte à decisão na agricultura.

Outro aspecto importante da OnTop é a representação de pragas e doenças, modelada pela Classe ***PestsAndDiseases***. Essa classe permite a identificação e o registro de ocorrências que podem afetar a produção, bem como a associação com ações de controle específicas. Isso é fundamental para o desenvolvimento de sistemas de suporte à decisão que auxiliem os agricultores na gestão de riscos fitossanitários.

Na construção de OnTop, os princípios de Gruber foram aplicados. A clareza foi assegurada por meio de definições precisas e formalizadas de cada conceito, evitando ambiguidades e garantindo que diferentes agentes possam interpretar a ontologia de maneira consistente. A coerência interna foi verificada utilizando *reasoners* compatíveis com OWL, que permitiram a identificação e correção de possíveis inconsistências lógicas.

A extensibilidade foi promovida pela adoção de uma estrutura modular, na qual classes genéricas podem ser especializadas para atender a necessidades específicas. Por exemplo, a Classe ***ManagementAction*** pode ser expandida para incluir novas práticas de manejo que venham a ser adotadas com o avanço tecnológico ou mudanças nas técnicas agrícolas. Já a Classe ***SoilCondition*** (Figura 4.7) pode ser expandida para trazer maior granularidade e especificidade aos tipos de solo, conforme necessidade.

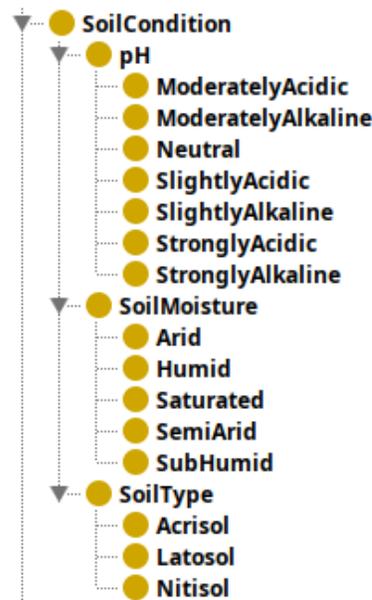


Figura 4.7 – Classe ***SoilCondition***, passível de expansão

Para minimizar o viés de codificação, OnTop foi especificada em um nível de abstração que evita dependências de implementações específicas ou formatos proprietários.

Quantidades e unidades de medida são representadas de forma genérica, permitindo a adaptação a sistemas de unidades utilizados em diferentes países ou regiões.

O comprometimento ontológico foi mantido ao mínimo necessário para representar adequadamente o domínio da produção de tabaco. Isso permite que a ontologia seja adotada por diferentes sistemas e agentes, sem impor suposições excessivas sobre o mundo.

OnTop incorpora relações temporais e de causalidade entre as classes, refletindo a natureza dinâmica do ciclo produtivo. Por exemplo, ações de manejo específicas são desencadeadas por condições climáticas ou pela detecção de pragas, e essas relações são modeladas na ontologia para permitir inferências sobre o momento adequado para realizar determinadas práticas.

A integração de fatores ambientais, como dados climáticos e características do solo, permite que OnTop seja utilizada em conjunto com sistemas de informação geográfica (GIS) e ferramentas de agricultura de precisão. Isso abre possibilidades para o desenvolvimento de sistemas de recomendação que forneçam orientações personalizadas aos agricultores, considerando as condições específicas de suas propriedades.

Em suma, OnTop representa uma contribuição para a representação formal do conhecimento no domínio da produção de tabaco. Sua construção fundamentada nos princípios de Gruber assegura que a ontologia seja um artefato robusto, capaz de promover o compartilhamento de conhecimento, a interoperabilidade entre sistemas e o suporte à tomada de decisão informada.

## 4.2 Avaliação de OnTop

Para avaliar a OnTop, utilizamos o OntoMetrics<sup>1</sup> (LANTOW, 2016). A análise de OnTop é mostrada na Figura 4.8.

OnTop apresenta considerável complexidade e expressividade, com um total de 1.291 axiomas, dos quais 880 são axiomas lógicos. Esses números refletem a capacidade de OnTop de formalizar uma ampla gama de relações e entidades envolvidas na produção de tabaco. Quanto maior o número de axiomas lógicos, maior é a capacidade da ontologia de capturar e representar nuances complexas do domínio modelado. Um valor alto indica que a ontologia é bastante detalhada e capaz de suportar raciocínios sofisticados.

OnTop possui 143 classes, oferecendo um arcabouço abrangente para representar as várias entidades e relações inerentes à produção de tabaco. Isso significa que

---

<sup>1</sup> OntoMetrics é uma ferramenta online que valida e exibe estatísticas sobre uma ontologia. Link: <https://ontometrics.informatik.uni-rostock.de/ontologymetrics/>

ontometrics.informatik.uni-rostock.de/ontologymetrics/ServletController

**OntologyID:** Optional.of(<http://www.semanticweb.org/luis-alves/ontologies/2024/8/tabak>)

**Base metrics** [more details](#) [Hide](#)

Axioms:	1291
Logical axioms count:	880
Class count:	143
Total classes count:	143
Object property count:	96
Total object properties count:	96
Data property count:	69
Total data properties count:	69
Properties count:	165
Individual count:	43
Total individuals count:	43
DL expressivity:	SROIF(D)

Figura 4.8 – Análise de OnTop por meio de OntoMetrics

a ontologia cobre uma vasta gama de conceitos específicos, desde condições climáticas e do solo até ações de manejo e estágios da produção de tabaco. Algumas classes primitivas, e.g., *ManagementAction*, alcançaram altos níveis de detalhamento, como pode ser visto na Figura 4.9.

A expressividade da lógica descritiva (DL) de OnTop é SROIF(D)<sup>2</sup>, o que implica suporte para hierarquias de papéis, inversões de papéis, transitividade e funcio-

<sup>2</sup> SROIF(D) é um nível de expressividade da Lógica Descritiva (DL), usado para representar e raciocinar sobre conhecimentos em ontologias. Cada letra indica uma capacidade: **S** (Hierarquias de papéis) permite organizar relações hierárquicas entre papéis, com herança de propriedades; **R** (Inclusão e Reflexividade de papéis) permite axiomas de inclusão de papéis e papéis reflexivos, relacionando um indivíduo consigo mesmo; **O** (Nominiais) possibilita o uso de indivíduos específicos como parte de definições de classes; **I** (Papéis Inversos) suporta raciocínio em ambos os sentidos de uma relação (por exemplo, "A é pai de B" implica "B é filho de A"); **F** (Propriedades Funcionais) restringe propriedades para terem apenas um valor por indivíduo (e.g., "Cada pessoa tem exatamente uma data de nascimento"); **D** (Tipos de Dados) permite raciocínio sobre tipos de dados (números inteiros, strings, etc.) e suas restrições. Esse nível de expressividade corresponde a um subconjunto do padrão OWL 2 DL, equilibrando expressividade e eficiência computacional.

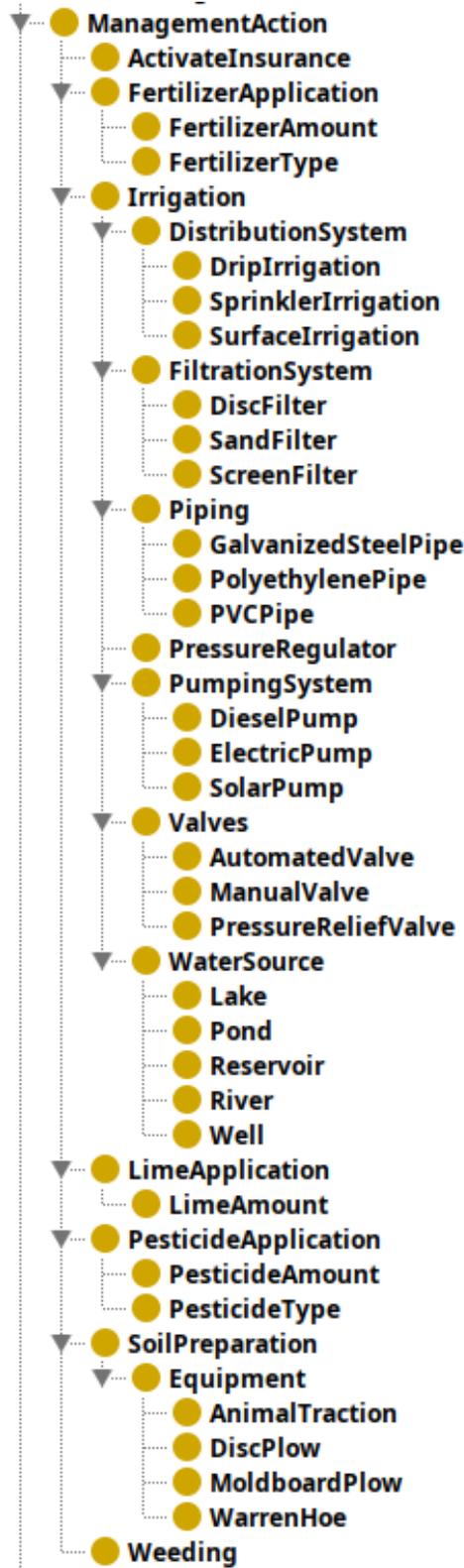


Figura 4.9 – Exemplo do nível de detalhamento da Classe *ManagementAction*.

nalidade. Em termos práticos, isso significa que a ontologia pode lidar com relações complexas e definir com precisão restrições sobre como diferentes entidades podem interagir.

Quanto mais alta a expressividade DL, maior a capacidade da ontologia de representar informações detalhadas, mas também maior a complexidade de processamento durante o raciocínio automatizado.

A riqueza do OnTop pode ser observada pelas quantidades de propriedades de objeto (96) e de dados (69), resultando em um total de 165 propriedades. Isso mostra que a ontologia possui uma representação diversificada de relações (propriedades de objeto) e atributos (propriedades de dados) relevantes para a produção de tabaco. Um número elevado de propriedades de objeto e dados sugere uma ontologia robusta e detalhada, permitindo uma modelagem precisa das interações e características de cada classe. Exemplos dessas propriedades de objeto (a) e de dados (b) são mostrados na Figura 4.10.

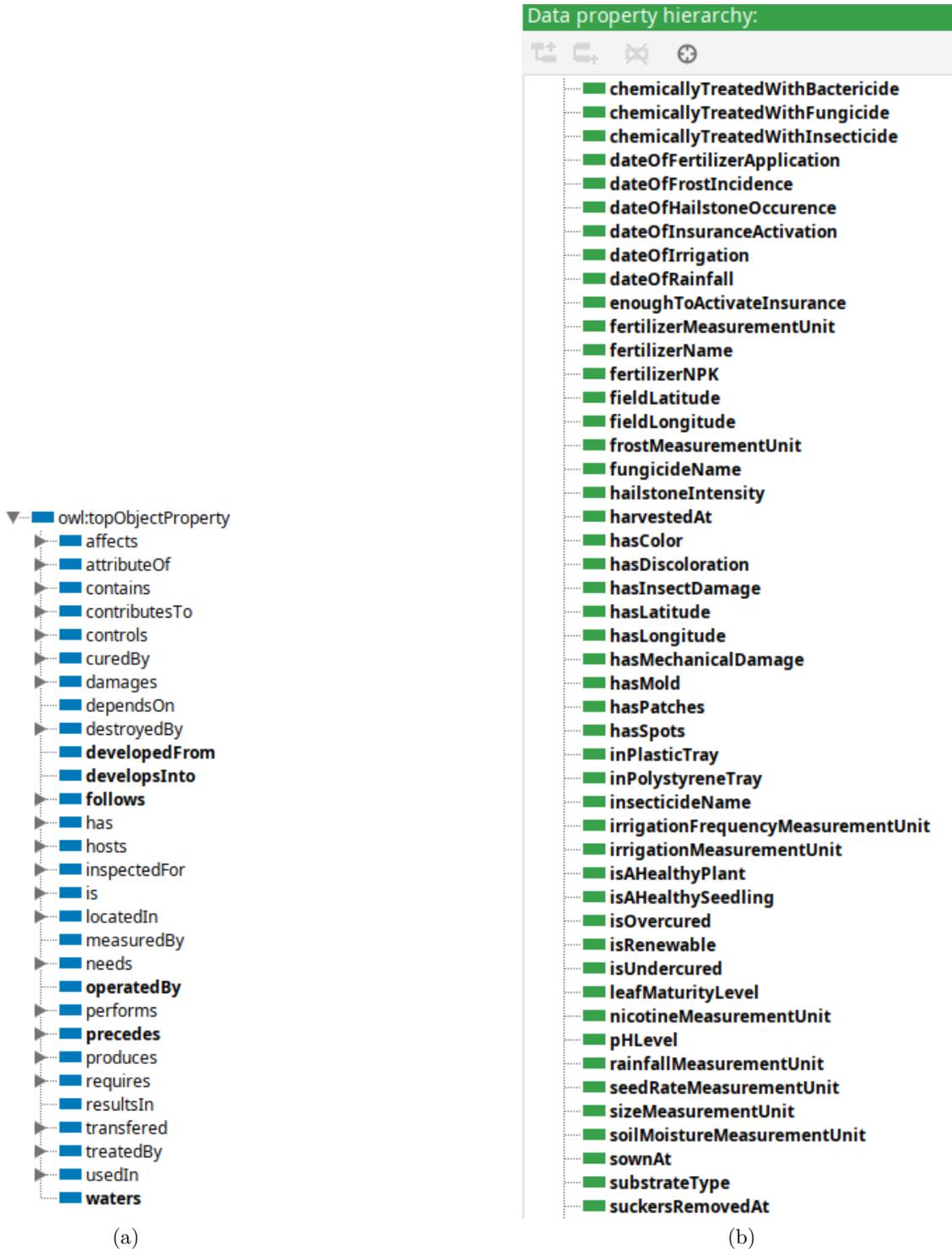


Figura 4.10 – Exemplos de propriedades de OnTop.

OnTop contém 43 indivíduos (Figura 4.11), que são instâncias iniciais da base de conhecimento. Esses indivíduos são exemplos concretos dos conceitos e relações definidos na ontologia, fornecendo um ponto de partida para futuras expansões do conhecimento. Em geral, quanto mais indivíduos, mais exemplos específicos e contextualizados a ontologia pode fornecer, facilitando a validação e aplicação prática do modelo. Um detalhamento de dois indivíduos (*seed\_sowing\_01* e *seed\_treatment\_01*) é mostrado na Figura 4.12.

Do ponto de vista estrutural, a ontologia contém 137 axiomas de subclasse (*SubClassOf*) e 11 axiomas de classe equivalente (*EquivalentClass*), o que reflete uma hierarquia bem organizada e a definição de classes sinônimas. A existência de 17 propriedades de objeto transitivas e 13 propriedades de objeto inversas aumenta a capacidade de raciocínio da ontologia, permitindo inferências em diferentes níveis da hierarquia de classes. Quanto maior o número de propriedades transitivas e inversas, mais complexas e inter-relacionadas podem ser as inferências, resultando em um modelo mais rico e realista.

A presença de 101 axiomas de domínio de propriedade de objeto e 100 axiomas de intervalo de propriedade de objeto assegura que as relações estão bem definidas no contexto de seus respectivos domínios e intervalos. Além disso, os 16 axiomas de propriedade de dados funcionais propostos garantem que certas propriedades são limitadas a valores únicos, adicionando mais uma camada de precisão ao modelo. Por exemplo, propriedades funcionais são úteis quando cada instância de uma classe pode ter apenas um valor para um dado atributo, como o número de identificação de um campo de tabaco.

As métricas do esquema sugerem *insights* adicionais sobre a riqueza da ontologia. Uma riqueza de herança de 0,958 e uma riqueza de relação de 0,460 indicam que o OnTop captura uma quantidade considerável de conhecimento hierárquico, com um alto grau de inter-relações entre as classes. Valores altos nesses parâmetros sugerem que a ontologia é capaz de representar relações complexas entre conceitos e que há uma estrutura hierárquica densa.

Todas as principais classes e subclasses estão anotadas com indicações da natureza daquela estrutura. Uma dessas anotações é mostrada na Figura 4.13.

A razão axioma-para-classe de 9,03 mostra um nível moderadamente alto de detalhamento na axiomatização de cada classe, refletindo a profundidade da ontologia na representação do conhecimento sobre a produção de tabaco. Quanto maior a razão axioma-para-classe, mais informações e restrições estão associadas a cada classe, o que pode indicar um modelo mais detalhado e rigoroso.

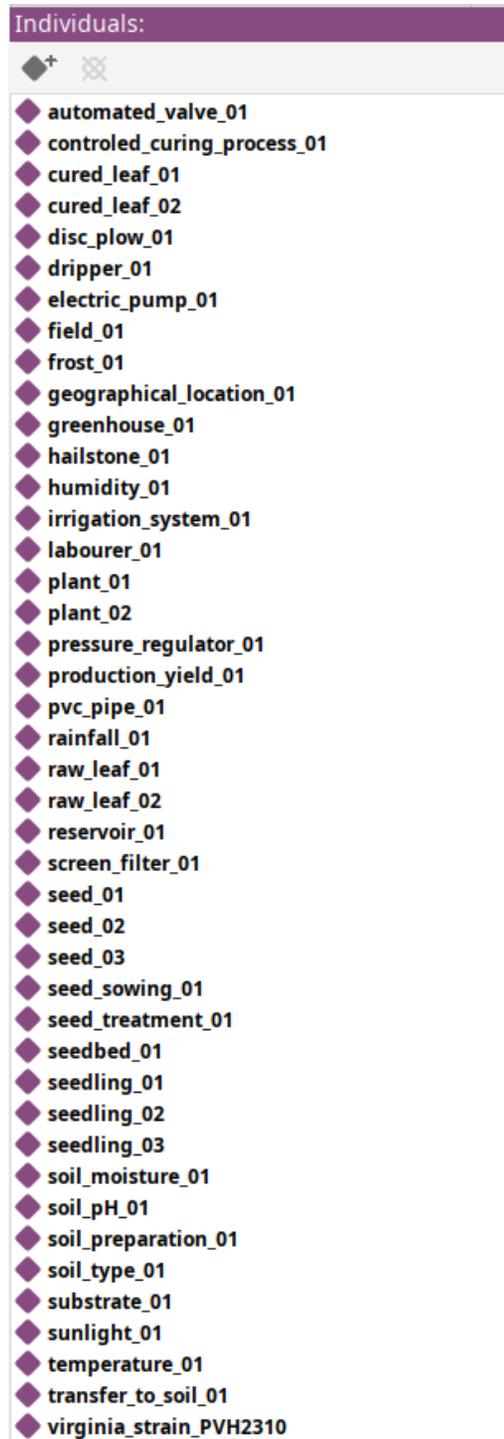


Figura 4.11 – Indivíduos declarados em OnTop

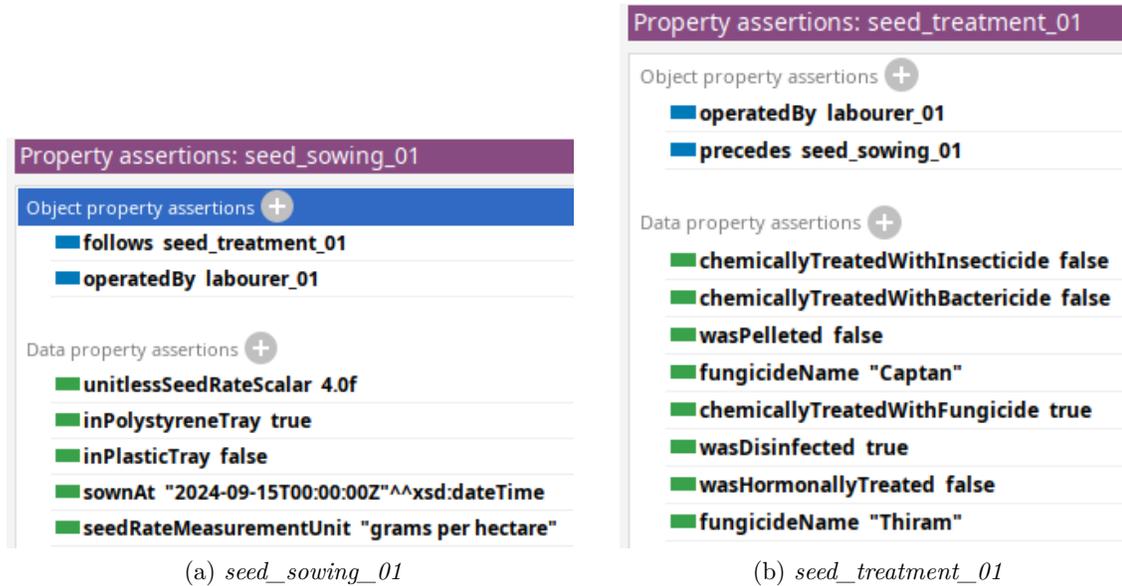


Figura 4.12 – Exemplos de indivíduos relacionados a *TobaccoSeedbed*.



Figura 4.13 – Exemplo de anotação na Classe *ClimateCondition*.

## 5 Aplicando OnTop em um Modelo de Linguagem Extenso (LLM)

A integração de ontologias com Modelos de Linguagem Extensos representa uma fronteira promissora no campo da inteligência artificial. LLMs (e.g., GPT-3 e seus sucessores) têm demonstrado capacidades notáveis em compreensão e geração de linguagem natural (BROWN *et al.*, 2020). No entanto, esses modelos enfrentam desafios relacionados à consistência, precisão factual e entendimento profundo de domínios específicos (BOMMASANI *et al.*, 2021). A incorporação de ontologias, como a OnTop, pode enriquecer o conhecimento de LLMs, proporcionando estrutura e semântica que auxiliam na geração de respostas mais precisas e contextualmente relevantes.

Neste capítulo, exploramos a aplicação da Ontologia da Produção de Tabaco (OnTop) em um modelo de linguagem extenso. Discutimos a metodologia empregada para integrar a ontologia ao LLM, os experimentos realizados e os resultados obtidos, destacando os benefícios e desafios dessa abordagem.

OnTop foi integrada a um LLM para avaliar sua eficácia na melhoria das respostas do modelo. A ontologia, desenvolvida em OWL, foi diretamente utilizada pelo LLM, sem a necessidade de conversão para outros formatos. O arquivo `.owl` contendo a OnTop foi disponibilizado ao LLM, permitindo que o modelo acessasse e interpretasse os conceitos e relações definidos na ontologia.

O objetivo principal desse experimento foi verificar se a inclusão de OnTop poderia enriquecer o conhecimento do LLM sobre o domínio da produção de tabaco, levando a respostas mais informadas e precisas. O LLM, ao ser alimentado com a ontologia, passa a ter acesso a uma representação formal e estruturada do conhecimento sobre o cultivo do tabaco, incluindo etapas do processo produtivo, condições climáticas e de solo, práticas de manejo, tipos de tabaco, entre outros.

Para este experimento, utilizamos o ChatGPT-4o<sup>1</sup> como modelo de linguagem em teste (UT-LLM). Inicialmente, foram feitas duas perguntas ao UT-LLM sobre a produção de tabaco, e as respostas foram registradas. Em seguida, enviamos ao UT-LLM o arquivo OWL contendo a OnTop e solicitamos que, considerando o arquivo antes de responder, respondesse às mesmas duas perguntas. As respostas foram analisadas para verificar se e em que medida a ontologia influenciou as respostas do UT-LLM. Perguntas

---

<sup>1</sup> ChatGPT é uma ferramenta de IA desenvolvida pela empresa OpenAI que fornece funcionalidades de chatbot. GPT significa *Generative Pre-trained Transformer*, um tipo de LLM.

e respostas estão em Inglês devido à ontologia ter sido construída neste idioma.

As perguntas (Q) feitas ao UT-LLM foram as seguintes:

**Q1)** *How does the treatment of seeds occur in tobacco production, and what is its influence?*

**Q2)** *How should the harvesting of tobacco work regarding the phases of it and actions between these phases?*

Como o UT-LLM tende a ser verboso, omitimos partes das respostas para facilitar a compreensão do seu raciocínio. Os principais pontos das respostas para cada pergunta (An) foram (destaques nossos):

**A1)** *[...] The seeds are extremely small, so **they are often coated to improve handling and sowing**. Key steps include: Seed Coating: Seeds are often coated with materials to make them easier to plant and manage. This helps in even distribution during sowing. Pre-germination Treatments: Seeds **may undergo pre-germination treatments**, including soaking in water or applying fungicides to prevent diseases. Controlled Environment for Germination: Tobacco seeds are typically **germinated in nursery beds or greenhouses**, where temperature, humidity, and light are carefully controlled to ensure uniform growth.*

**A2)** *[...] Tobacco harvesting occurs in multiple phases depending on the type of tobacco and its intended use. The process generally follows these stages: Priming (Bottom Leaves): The first phase involves harvesting the bottom leaves, which are typically riper earlier. These leaves are crucial for certain types of tobacco products, like cigarettes. Farmers harvest them manually by pulling off the mature leaves. Mid-Stalk Leaves: Once the bottom leaves are harvested, attention turns to the middle stalk leaves, which are typically more valuable due to their higher nicotine content and better quality for curing. Top Leaves: The final phase is the harvesting of the top leaves, which mature last. These leaves are often thicker and contain more nicotine. Actions Between Phases: Curing: **After each phase, the leaves undergo curing** to preserve and enhance the flavor and aroma of the tobacco. Curing methods include air curing, flue curing, or sun curing, depending on the type of tobacco. Proper curing between harvest phases is vital for leaf quality. [...]*

Na segunda parte do teste, enviamos ao UT-LLM o arquivo OWL da OnTop com a seguinte mensagem: "*Now, read integrally this ontology in the OWL file I am sending attached, and answer the same questions (Q1 and Q2)*".

Os principais pontos das respostas enriquecidas (Aen) foram (destaques nossos):

*Ae1) [...] The treatment process includes **not only coating** but also protecting the seeds against potential threats such as insect and bacterial damage. [...] Treatments may involve **insect control measures to ensure healthy seedlings**. Irrigation Systems: Once the seeds have germinated, **irrigation becomes a key factor**. [...] Emphasizing the importance of maintaining proper water levels to support healthy seedling growth. Seedling Health Monitoring: Plant health is monitored throughout the early stages, ensuring that treatments are adjusted based on factors like **irrigation, pest control, and soil conditions**. Thus, the treatment of seeds not only involves germination but also an ongoing system of health monitoring and environmental management, with a focus on protecting against pests and ensuring optimal growing conditions.*

*Ae2) [...] The first phase of harvesting is the picking of the bottom leaves, which mature earlier and are manually harvested. **Topping is another phase which refers to removing the flower from the plant**. This helps direct more energy to the leaves, enhancing their quality. [...] Actions Between Phases: Curing: Curing is a critical action between harvest phases, and additional focus must be applied on proper techniques to avoid issues **like mold formation on the leaves**. This reinforces the importance of careful environmental control during curing. [...] Also important is a system of ongoing health monitoring during each phase, ensuring that any issues like **insect damage or under-curing** are addressed early.*

Ao comparar as respostas antes e depois da integração da OnTop, observamos que o UT-LLM passou a incluir informações mais detalhadas e específicas sobre a produção de tabaco.

## 6 Discussão sobre a Aplicação

Neste capítulo, analisamos os resultados obtidos a partir do experimento de aplicação da Ontologia de Produção de Tabaco (OnTop) em um Modelo de Linguagem Extenso (LLM). Compararemos as respostas fornecidas pelo LLM antes e depois de ter acesso à OnTop, avaliando como a ontologia influenciou a qualidade e a precisão das respostas.

Consideramos as respostas dadas pelo LLM antes de conhecer a OnTop (A1 e A2) como razoáveis. Na resposta à primeira pergunta (**A1**), o modelo apresentou como ocorre o tratamento de sementes, incluindo exemplos como o recobrimento das sementes para melhorar o manuseio e a semeadura, e tratamentos pré-germinação, como a imersão em água. Embora a resposta careça de detalhes, é em geral confiável.

Na resposta à segunda pergunta (**A2**), o LLM apresentou uma visão simplificada do processo de colheita do tabaco, não refletindo a rigidez do cronograma das fases. Ao comentar sobre as ações entre as fases, o modelo forneceu uma resposta incorreta, afirmando que "*after each phase, the leaves undergo curing*". Isso é incorreto, pois a cura não ocorre entre as fases de colheita, mas após a colheita completa.

Após a interação com a OnTop, observamos mudanças significativas nas respostas do LLM. Na resposta à primeira pergunta (**Ae1**), o modelo mencionou não apenas o recobrimento das sementes, mas também a proteção contra ameaças como insetos e bactérias, medidas de controle de insetos e a importância da irrigação após a germinação. Além disso, destacou o monitoramento da saúde das mudas, ajustando tratamentos com base em fatores como irrigação, controle de pragas e condições do solo.

Essas informações adicionais não estavam presentes na resposta inicial (**A1**). O LLM passou a adotar uma abordagem mais global do que na primeira resposta, enfatizando a importância da irrigação, controle de pragas e condições do solo para que os tratamentos de sementes tenham o efeito desejado. Isso indica que o conhecimento da OnTop permitiu ao modelo enriquecer sua resposta, incorporando conceitos críticos presentes na ontologia.

Na resposta à segunda pergunta (**Ae2**), o LLM ainda simplifica o processo, mas adiciona outra camada de complexidade ao citar o processo de *capação* (*topping* em inglês), que estava ausente em **A2**. A *capação* refere-se à remoção da flor da planta, direcionando mais energia e nutrientes para as folhas e melhorando sua qualidade. Sobre as ações entre as fases, o modelo ainda apresenta a afirmação incorreta de que "*curing is a critical action between harvest phases*", o que é um ponto negativo. No entanto, reconhece

a importância de evitar a formação de mofo (um problema significativo na cura do tabaco, não mencionado em **A2**) e o monitoramento da saúde para prevenir danos por insetos e subcura, ambos conceitos presentes na OnTop.

Em resumo, o LLM utilizou o conhecimento da OnTop para melhorar suas respostas, adicionando conceitos críticos e fornecendo uma visão mais ampla de todo o ciclo de vida do tabaco, embora tenha cometido um erro ao sugerir a cura entre as fases de colheita. Este erro indica que, apesar do enriquecimento proporcionado pela ontologia, ainda existem limitações na interpretação completa e precisa de todos os conceitos pelo LLM.

A partir desses resultados, conclui-se que a aplicação da OnTop em um LLM pode enriquecer significativamente a capacidade do modelo de fornecer orientações corretas aos usuários sobre a produção de tabaco. A estrutura ontológica da OnTop forneceu ao LLM um arcabouço de conhecimento especializado, que foi refletido nas respostas aprimoradas. O modelo passou a incluir detalhes importantes sobre tratamentos de sementes, práticas de manejo e considerações críticas durante a colheita e pós-colheita.

Este experimento sugere o potencial da integração de ontologias com modelos de linguagem para melhorar a qualidade das interações em domínios específicos. Ao fornecer ao LLM acesso a uma ontologia bem estruturada e abrangente como a OnTop, é possível superar limitações de conhecimento e aumentar a precisão e relevância das respostas geradas. No entanto, também evidencia a necessidade de melhorias na forma como os LLMs interpretam e aplicam o conhecimento ontológico, para evitar erros como o identificado na resposta sobre a cura entre fases de colheita.

Após a integração da OnTop com o LLM as respostas foram mais detalhadas e contextualizadas, sugerindo que ontologias podem ser ferramentas valiosas para enriquecer o desempenho de modelos de linguagem em domínios especializados. Este avanço tem implicações positivas para o desenvolvimento de sistemas de recomendação e de suporte à decisão na agricultura e em outros setores que se beneficiam de conhecimento especializado estruturado.

Por fim, para validar a análise feita e tirar proveito das capacidades de língua natural da LLM conduzimos um pequeno teste. Enviamos o texto do Capítulo 5 para a LLM (note-se que esse capítulo não tem a análise dos autores acerca das respostas) com a seguinte instrução: *According to the text above, how would you say the LLM changed the answer? Did it alter the response in any significant way? Did the ontology cause a change in the quality and completeness of the answer?*. A resposta do ChatGPT-4o foi a seguinte (resposta essa, de forma geral, muito próxima da avaliação dos autores):

## Resposta do ChatGPT-4o

Based on the analysis provided, the inclusion of the Ontology of Tobacco Production (OnTop) in the LLM had a notable impact on the quality and completeness of the answers:

### Key Changes in the Responses

#### 1. Increased Specificity and Domain Knowledge:

- **Before Ontology Integration:** The LLM's responses were more general, focusing on standard agricultural processes like seed coating, pre-germination treatments, and curing techniques.
- **After Ontology Integration:** The responses incorporated more specific and nuanced domain knowledge, such as:
  - Insect control measures during seed treatment.
  - Importance of irrigation systems and environmental factors for seedling growth.
  - Reference to "topping" as a specific phase in tobacco harvesting.
  - Detailed discussion on preventing mold during curing and addressing under-curing or insect damage.

#### 2. Enhanced Contextual Relevance:

- With the ontology, the LLM provided richer context by connecting processes like health monitoring to broader environmental management practices, reflecting a deeper understanding of the domain.

#### 3. Addition of New Concepts:

- The ontology introduced previously absent concepts such as:
  - The role of irrigation in seedling development.
  - "Topping" during harvesting to enhance leaf quality.
  - Focused techniques for curing, like mold prevention and addressing specific curing issues.

### Impact on Quality and Completeness

- **Quality:** The responses became more detailed and precise, with a higher level of factual accuracy and alignment with domain-specific practices.
- **Completeness:** The ontology enriched the answers by filling gaps and providing a more comprehensive explanation of processes.

### Conclusion

The ontology caused a **significant improvement** in the answers' quality and completeness. It enhanced the LLM's ability to deliver domain-specific knowledge that was more aligned with practical realities. By integrating structured ontological knowledge, the LLM could reason more effectively about the domain, resulting in responses that were not only more informative but also contextually relevant. This demonstrates the potential of ontologies to improve LLM outputs, particularly in specialized fields.

## 7 Conclusão

Ao longo desta dissertação, exploramos a importância e o impacto da aplicação de ontologias no domínio agrícola, com foco específico na produção de tabaco. A Ontologia da Produção de Tabaco (OnTop) foi desenvolvida como uma proposta inovadora para formalizar e representar de maneira estruturada todo o ciclo de vida do cultivo de tabaco. Fundamentada em princípios tradicionais de design de ontologias, OnTop visa a promover compartilhamento de conhecimento, interoperabilidade entre sistemas e suporte à recomendação e à tomada de decisão no setor agrícola.

Foram discutidos os conceitos fundamentais e critérios para a construção de ontologias voltadas ao compartilhamento de conhecimento. Em seguida, realizamos uma revisão sistemática da literatura sobre sistemas de recomendação baseados em GIS na agricultura, identificando lacunas e oportunidades para a aplicação de ontologias de domínio específicas. OnTop foi então apresentada em detalhes, incluindo exemplos de sua formalização em Lógica Descritiva e sintaxe Manchester.

A aplicação de OnTop em um Modelo de Linguagem Extenso (LLM) sugeriu o potencial da integração de ontologias com modelos de linguagem para melhorar a qualidade das interações em domínios especializados. O experimento realizado evidenciou que o LLM sob teste, ao ter acesso à OnTop, forneceu respostas mais precisas e contextualizadas sobre a produção de tabaco, incorporando conceitos críticos e proporcionando uma visão mais abrangente do ciclo produtivo.

### 7.1 Resultados da Pesquisa

A pesquisa realizada contribuiu para o campo da representação formal do conhecimento na agricultura, apresentando os seguintes resultados:

- **Desenvolvimento da OnTop:** Foi criada a primeira ontologia formal dedicada à produção de tabaco, estruturando de maneira sistemática os conceitos, relações e propriedades relevantes do domínio. A OnTop permite a representação detalhada de etapas do ciclo produtivo, práticas de manejo, condições climáticas e de solo, entre outras.
- **Aplicação dos Princípios de Gruber:** A OnTop foi desenvolvida seguindo os princípios de design de ontologias propostos por Gruber (GRUBER, 1995), assegurando clareza, coerência, extensibilidade, mínimo viés de codificação e mínimo

comprometimento ontológico. Esse é, assim, um exemplo da utilização dos princípios teóricos de Gruber em uma aplicação para o mundo real.

- **Integração com Modelos de Linguagem:** O experimento de aplicação da OnTop em um LLM sugeriu que a ontologia pode enriquecer a capacidade do modelo de fornecer orientações corretas e detalhadas sobre a produção de tabaco. As respostas do LLM passaram a incluir informações mais precisas e a refletir um entendimento mais profundo do domínio, evidenciando o potencial da integração entre ontologias e modelos de linguagem.
- **Contribuição para a Agricultura de Precisão:** A OnTop pode estabelecer uma base para o desenvolvimento de sistemas de suporte à decisão, sistemas de recomendação e outras ferramentas tecnológicas na agricultura. Ao representar formalmente o conhecimento do domínio, a ontologia facilita a integração de dados heterogêneos e marginais.
- **Modelo para Outras Culturas:** A metodologia e os princípios aplicados no desenvolvimento da OnTop podem servir de referência para a criação de ontologias em outros domínios agrícolas, ampliando o impacto desta pesquisa para além da produção de tabaco.

## 7.2 Limitações e Trabalhos Futuros

Apesar dos avanços alcançados, a pesquisa apresenta algumas limitações que abrem oportunidades para trabalhos futuros:

- **Detalhamento Específico:** Embora a OnTop inclua os principais conceitos e relações da produção de tabaco, há espaço para aprofundar detalhes específicos. Por exemplo, incorporar informações sobre o microbioma do solo, que influencia a saúde das plantas e a qualidade do tabaco produzido.
- **Capacidades Temporais:** A ontologia poderia ser ampliada para melhor representar aspectos temporais do ciclo produtivo, como a duração de cada etapa, sazonalidade e dependências temporais entre práticas de manejo. Isso permitiria modelar cronogramas de cultivo com maior precisão.
- **Integração de Dados em Tempo Real:** Futuras pesquisas podem explorar a integração da OnTop com sistemas que fornecem dados em tempo real, como sensores de campo, estações meteorológicas e imagens de satélite. Essa integração potencializaria a capacidade de monitoramento e resposta rápida a eventos críticos.

- **Validação Empírica Ampliada:** Embora o experimento com o LLM tenha mostrado resultados promissores, seria valioso realizar validações adicionais com outros modelos de linguagem e em diferentes contextos, incluindo testes com indústrias do setor do tabaco para avaliar a usabilidade e a eficácia prática da ontologia.
- **Interoperabilidade com Outras Ontologias:** Explorar a integração da OnTop com ontologias relacionadas, como aquelas voltadas para a gestão de recursos hídricos ou sistemas de comercialização, poderia ampliar o escopo e a utilidade da ontologia. Ademais, uma análise mais profunda é requerida para conectar o domínio proposto a uma ontologia de topo (e.g., BFO (SMITH *et al.*, 2006) ou UFO (GUIZZARDI *et al.*, 2013)).
- **Considerações Socioeconômicas:** Incorporar aspectos socioeconômicos na ontologia, como políticas agrícolas, regulamentações e restrições, forneceria uma visão mais holística do domínio e apoiaria tomadas de decisão em níveis estratégicos.
- **Automação de Processos Agrícolas:** Investigar como a OnTop pode ser utilizada em conjunto com tecnologias de automação e robótica na agricultura, contribuindo para o desenvolvimento de sistemas autônomos de manejo e colheita.

Em suma, a Ontologia de Produção de Tabaco representa um passo significativo em direção à representação formal do conhecimento agrícola. Ao abordar as limitações atuais e explorar as direções sugeridas para trabalhos futuros, espera-se que OnTop continue evoluindo e contribuindo para a inovação e a modernização na agricultura, especialmente na cultura do tabaco.

### 7.3 Considerações éticas

Estamos plenamente cientes dos efeitos negativos amplamente documentados do tabagismo para a saúde pública e para a sociedade, incluindo suas contribuições para doenças, dependência e impactos econômicos associados. Contudo, esta dissertação não tem como objetivo promover o consumo de tabaco. Em vez disso, ela aborda a produção agrícola do tabaco como uma realidade socioeconômica.

O tabaco continua sendo uma cultura essencial em diversas regiões, gerando o sustento de milhões de pessoas (agricultores, trabalhadores e suas famílias). Pesquisas voltadas à melhoria da produção de tabaco não representam um endosso ao seu consumo final, mas sim o reconhecimento da necessidade de apoiar as comunidades cuja subsistência depende dessa cultura. Ao avançar em práticas agrícolas, reduzir desperdícios e melhorar a sustentabilidade, esta pesquisa contribui para o bem-estar dos trabalhadores e de suas

famílias, promovendo práticas éticas e eficientes em um setor que continuará a existir independentemente da realização de pesquisas acadêmicas.

Os resultados e as metodologias apresentados também têm o potencial de inspirar aplicações mais amplas na agricultura de precisão, beneficiando culturas além do tabaco. Isso destaca o papel da ciência agrícola na promoção de práticas sustentáveis e na melhoria das condições de vida em múltiplos setores.

## 7.4 Resultados Técnico-científicos

**Artigo 1: Revisão de Literatura.** Artigo científico de revisão sistemática de literatura, submetido ao journal *Cadernos de Ciência e Tecnologia (CC&T)* da EMBRAPA, com o título "*GIS-based Recommender Systems for Agriculture: a systematic literature review*". O objetivo deste artigo foi apresentar o estado da arte sobre o uso de Sistemas de Recomendação (RS) baseados em Sistemas de Informação Geográfica (GIS) aplicados à agricultura. Através de uma revisão sistemática da literatura, buscou-se descrever e classificar os estudos realizados nos últimos seis anos, analisando as principais tendências, abordagens metodológicas, contribuições e limitações das pesquisas no campo. Além disso, o artigo teve como propósito identificar as lacunas existentes na literatura e destacar os desafios de pesquisa que precisam ser enfrentados para o desenvolvimento futuro de auxiliares de tomada de decisão que possam lidar de forma eficiente com os diversos fatores envolvidos na agricultura, como clima, solo e mecanização.

**Artigo 2: Principal.** Artigo científico de construção da OnTop, submetido ao journal *Revista de Informática Teórica e Aplicada*, com o título "*An Ontology of Tobacco Production: enriching LLM-based decision support*". O objetivo deste artigo foi apresentar a Ontologia da Produção de Tabaco (OnTop), que foi projetada para auxiliar na otimização das práticas de manejo de culturas de tabaco, integrando fatores ambientais em um modelo formal. OnTop é pioneira ao formalizar o ciclo de vida da produção de tabaco em conjunto com dados de solo e clima. Este trabalho define e aplica a OnTop para enriquecer um Modelo de Linguagem Extenso (LLM), permitindo suporte a decisões baseado em ontologia. As principais contribuições incluem: i) uma ontologia de domínio que formaliza o conhecimento sobre a produção de tabaco orientada por dados; e ii) um framework consultável, LLM enriquecida, que permite aos especialistas obter respostas complexas sobre agronomia.

**Ontologia: OWL de OnTop.** Ontologia da Produção de Tabaco, de livre acesso, disponível no repositório Github (ALVES; ROSA, 2024).

## Referências

- AGROPECUÁRIA, F. P. da. *Cultura do tabaco*. 2020. Disponível em: <<https://fpagropecuaria.org.br/2020/06/29/cultura-do-tabaco/?pdf=30788>>. Citado na página 11.
- ALVES, L. F. M.; BONACIN, R.; FRANCO, F. d. R. An ontology of tobacco production: enriching llm-based decision support. *Revista de Informática Teórica e Aplicada (submitted)*, 2024. Citado 2 vezes nas páginas 27 e 28.
- ALVES, L. F. M.; FRANCO, F. d. R. Gis-based recommender systems for agriculture: A systematic literature review. *Cadernos de Ciência & Tecnologia (submitted)*, 2024. Citado na página 20.
- ALVES, L. F. M.; ROSA, F. F. *OnTop ontology*. 2024. <<https://github.com/lfmalves/OnTop>>. Citado 2 vezes nas páginas 27 e 54.
- ANLEY, M. B.; TESEMA, T. B. A collaborative approach to build a kbs for crop selection: Combining experts knowledge and machine learning knowledge discovery. In: *Communications in Computer and Information Science*. [S.l.]: Springer International Publishing, 2019. p. 80–92. Citado 2 vezes nas páginas 21 e 22.
- AVEZBOYEV, S.; SHARIPOV, S.; XUJAKELDIEV, K. Development of projects for recultivation of lands using gis technologies. *IOP Conference Series: Earth and Environmental Science*, v. 1138, n. 1, p. 012019, 2023. Citado 2 vezes nas páginas 21 e 22.
- BOMMASANI, R. *et al.* On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2021. Citado na página 45.
- BROWN, T. B. *et al.* Language models are few-shot learners. *Advances in Neural Information Processing Systems*, v. 33, p. 1877–1901, 2020. Citado na página 45.
- CHEBOTAREV, P. M. *et al.* Assessment of the soil-resource potential of the territory based on the processing of remote sensing data in a gis environment. In: *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. [S.l.]: IEEE, 2020. Citado na página 22.
- COX, S.; WEST, R.; NOTLEY, C.; SOAR, K.; HASTINGS, J. Toward an ontology of tobacco, nicotine and vaping products. *Addiction*, v. 118, n. 1, p. 177–188, 2023. Citado 2 vezes nas páginas 24 e 25.
- DABRE, K. R.; LOPES, H. R.; DMONTE, S. S. Intelligent decision support system for smart agriculture. In: *2018 International Conference on Smart City and Emerging Technology (ICSCET)*. [S.l.]: IEEE, 2018. Citado na página 22.
- DOSHI, Z. *et al.* Agroconsultant: Intelligent crop recommendation system using machine learning algorithms. In: *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. [S.l.]: IEEE, 2018. Citado 2 vezes nas páginas 21 e 22.

- ED-DAOUDI, M. *et al.* Machine learning algorithms in agriculture: A review. *Journal of Agricultural Informatics*, 2023. Citado 3 vezes nas páginas 20, 21 e 22.
- EMATER-AL; EMBRATER; EMBRAPA. *Sistema de Produção para a Cultura do Fumo*. Maceió, 1981. Citado na página 17.
- EZRA, A. *et al.* Application of geographic information system (gis) and analytical hierarchy process (ahp) in suitability mapping of some selected cereal crops in northeastern nigeria. *BIMA Journal of Science and Technology*, v. 7, n. 01, p. 33–46, 2023. Citado 2 vezes nas páginas 21 e 22.
- FARHEEN, N. Rainfall prediction and suitable crop suggestion using machine learning prediction algorithms. In: *Advances in Intelligent Systems and Computing*. [S.l.]: Springer Nature Singapore, 2021. p. 497–513. Citado 2 vezes nas páginas 21 e 22.
- FEGADE, T. K.; PAWAR, B. V. Crop prediction using artificial neural network and support vector machine. In: *Data Management, Analytics and Innovation*. [S.l.]: Springer Singapore, 2020. p. 311–324. Citado na página 22.
- GAO, Z.; LI, Z.; CHEN, J. Planning and suggestions for the walnut production areas in beijing. In: *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*. [S.l.]: IEEE, 2018. Citado 2 vezes nas páginas 21 e 22.
- GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, v. 43, n. 5-6, p. 907–928, 1995. Citado 8 vezes nas páginas 11, 12, 13, 14, 24, 26, 27 e 51.
- GUIZZARDI, G.; WAGNER, G.; RECKER, J.; EVERMANN, J. Towards ontological foundations for the conceptual modeling of events. In: *Proceedings of the 32nd International Conference on Conceptual Modeling (ER 2013)*. [S.l.]: Springer, 2013. p. 327–341. Citado na página 53.
- HOFFMANN, J. C.; MARCON, G. *Tabaco - Manejo Conservacionista de Solo*. Curitiba: Serviço Nacional de Aprendizagem Rural (SENAR-PR), 2012. Citado na página 17.
- IYER, R. T. *et al.* Geostatistical modeling of soil nutrients and development of a mobile application for site-specific fertilizer recommendation. In: *2019 IEEE Recent Advances in Geoscience and Remote Sensing: Technologies, Standards and Applications (TENGARSS)*. [S.l.]: IEEE, 2019. Citado na página 22.
- KITCHENHAM, B. *Procedures for performing systematic reviews*. [S.l.], 2004. 28 p. (TR/SE-0401, v. 33). Citado na página 20.
- LANTOW, B. Ontometrics: Putting metrics into use for ontology evaluation. In: *International Conference on Knowledge Engineering and Ontology Development*. [s.n.], 2016. Disponível em: <<https://api.semanticscholar.org/CorpusID:42135769>>. Citado na página 37.
- LEONE, C. A. Genetics: Lysenko versus mendel. *Transactions of the Kansas Academy of Science*, v. 55, n. 4, p. 369–380, 1952. ISSN 0022-8443. Disponível em: <<https://www.jstor.org/stable/3625986>>. Citado na página 17.

- MUPANGWA, W. *et al.* Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in eastern and southern africa. *SN Applied Sciences*, v. 2, n. 5, 2020. Citado na página 22.
- MUSEN, M. A. The protégé project. *AI Matters*, v. 1, n. 4, p. 4–12, June 2015. ISSN 2372-3483. Citado na página 27.
- NOTLEY, C.; WEST, R.; SOAR, K.; HASTINGS, J. Toward an ontology of identity-related constructs in addiction, with examples from nicotine and tobacco research. *Addiction*, v. 118, n. 3, 2022. Citado na página 25.
- NUNGULA, E. Z. *et al.* Land degradation unmasked as the key constraint in sunflower (*helianthus annuus*) production: Role of gis in revitalizing this vital sector. *Cogent Food and Agriculture*, v. 9, n. 2, 2023. Citado 2 vezes nas páginas 21 e 22.
- OSMAN, H. *et al.* Ontology based recommendation system for predicting cultivation and harvesting timings using support vector regression. In: *Software Engineering Perspectives in Intelligent Systems*. Cham: Springer International Publishing, 2020. p. 813–824. Citado 2 vezes nas páginas 21 e 22.
- PETERS, D. P. C. *et al.* Ai recommender system with ml for agricultural research. *IT Professional*, v. 22, n. 3, p. 30–32, 2020. Citado 2 vezes nas páginas 21 e 22.
- PIERCE, F. J.; CLAY, D. E.; SCHULER, R. T. *Applications of GIS in precision agriculture*. [S.l.]: John Wiley and Sons, 2007. Citado 2 vezes nas páginas 11 e 20.
- PRASAD, N. R.; PATEL, N. R.; DANODIA, A. Crop yield prediction in cotton for regional level using random forest approach. *Spatial Information Research*, v. 29, p. 195–206, 2021. Citado 2 vezes nas páginas 21 e 22.
- PRIYADHARSHINI *et al.* Intelligent crop recommendation system using machine learning. In: *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. [S.l.]: IEEE, 2021. Citado 2 vezes nas páginas 21 e 22.
- RAHMAN, S. A. Z.; MITRA, K. C.; ISLAM, S. M. M. Soil classification using machine learning methods and crop suggestion based on soil series. In: *2018 21st International Conference of Computer and Information Technology (ICCIT)*. [S.l.]: IEEE, 2018. Citado 2 vezes nas páginas 21 e 22.
- RAJAMANI, K.; HARI, N.; RAJASHEKAR, M. Soil fertility evaluation and gps-gis-based soil nutrient mapping of krishi vigyan kendra, palem, telangana. *International Research Journal of Pure and Applied Chemistry*, p. 139–145, 2020. Citado na página 21.
- RAMCHANDANI, L.; PATEL, S.; SUTHAR, K. A multiple criteria-based context-aware recommendation system for agro-cloud. In: *Advances in Computing and Intelligent Systems*. [S.l.]: Springer Singapore, 2020. p. 557–564. Citado 2 vezes nas páginas 21 e 22.
- RIZALDI, T. *et al.* Decision support system for land selection to increase crops productivity in jember regency using learning vector quantization (lvq). In: *2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*. [S.l.]: IEEE, 2019. Citado 2 vezes nas páginas 21 e 22.

RODRÍGUEZ-GARCÍA, M. ; GARCÍA-SÁNCHEZ, F. Croppesto: An ontology model for identifying and managing plant pests and diseases. In: *Communications in Computer and Information Science*. [S.l.]: Springer International Publishing, 2020. p. 18–29. Citado 2 vezes nas páginas 21 e 22.

SMITH, B.; GRENON, P.; OBERLE, D. Basic formal ontology (bfo). In: OBERLE, D. (Ed.). *The Applied Ontology Handbook: An Introduction to Ontology in Information Systems*. [S.l.]: Springer, 2006. p. 47–68. Citado na página 53.

SUCHITHRA, M. S.; PAI, M. L. Data mining-based geospatial clustering for suitable recommendation system. In: *2020 International Conference on Inventive Computation Technologies (ICICT)*. [S.l.]: IEEE, 2020. Citado 2 vezes nas páginas 21 e 22.

TINOCO, J. *et al.* Artificial neural networks for rock and soil cutting slopes stability condition prediction. In: *Sustainable Civil Infrastructures*. Cham: Springer International Publishing, 2019. p. 105–114. Citado 2 vezes nas páginas 21 e 22.

YUDHANA, A. *et al.* Spatial distribution of soil nutrient content for sustainable rice agriculture using geographic information system and naïve bayes classifier. *International journal on smart sensing and intelligent systems*, v. 16, n. 1, 2023. Citado 2 vezes nas páginas 21 e 22.

YUSIANTO, R. *et al.* Spatial analysis for crop land suitability evaluation: A case study of potatoes cultivation in wonosobo, indonesia. In: *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*. [S.l.]: IEEE, 2020. Citado na página 22.

ZERMAS, D. *et al.* A methodology for the detection of nitrogen deficiency in corn fields using high-resolution rgb imagery. *IEEE transactions on automation science and engineering: a publication of the IEEE Robotics and Automation Society*, v. 18, n. 4, p. 1879–1891, 2021. Citado 2 vezes nas páginas 21 e 22.