An Ontology-Based Intelligent Interactive Knowledge Interface for Groundnut Crop Information

Purvi H. Bhensdadia, C. K. Bhensdadia
Computer Engineering Department,
Dharmsinh Desai University, Nadiad, 387001, Gujarat, India

***Abstract*—This paper presents an ontology-based interactive interface designed to provide farmers in Gujarat with information related to groundnut crops. An ontology specific to the groundnut crop was developed and used to create a semantic question-answering (QA) interface. The proposed QA interface converts natural language question into SPARQL Query and provides answer using the backbone ontology. The overall performance of the system is at par with the existing semantic QA system. Overall accuracy of QA System is 80%.**

Keywords*—*Agriculture ontology; ontology construction; question answer system; groundnut ontology

I. Introduction

In India, agriculture is of paramount importance. This plays a crucial role in the development of rural areas. Gujarat is a leading producer of cash crops like cotton and groundnut. An estimated 20 lakh hectares of groundnuts are farmed in Gujarat each year, with a total production of roughly 26 lakh tons.[1]

A significant amount of data is available in the agricultural sector in the form of text documents, spreadsheets, and tables. There are many websites which have groundnut crop data in a factual form such as Farmer's Portal[2], mKishan[3], and i-Khedut(for groundnut crop)[4]. In addition, some online application exist for groundnut crops, such as i-khedut[5], magfadi[6], Chhomasu magfadi ma pramanit bij[7], and Khedut mol[8]. These websites and applications cannot perform semantic searches or reasoning.

The primary drawback of the existing applications or systems is the dependency on agriculture experts or other educated farmers to answer farmers' queries. These web applications are frequently used by farmers to express natural language queries that are answered by agriculture experts [1]. However, the expert might not always be available to respond to all the farmers' queries, which can create a communication gap between the farmer and the agriculture expert. To bridge this gap, semantic search techniques [2] can be employed. Semantic search [3] enhances the search capability by understanding the context and intent behind the queries, thus providing more relevant and accurate results [4]. This approach can automate the process of answering farmers' queries, making it possible to access information without waiting for an expert's response. Implementing a semantic search system can significantly improve the efficiency and effectiveness of information retrieval in agriculture, ultimately benefiting farmers by providing timely and accurate information.

An ontology-based question-answer system has been developed to interpret farming-related queries and provide relevant suggestions. This system leverages the structured knowledge within the ontology to provide context-aware responses, bridging the gap between farmers and agricultural experts. By utilizing this approach, farmers can receive immediate and relevant answers to their questions, enhancing their ability to make informed decisions about their crops and farming practices. This innovation not only improves the accessibility of agricultural knowledge but also empowers farmers with the tools necessary for efficient and effective farming.

The major contribution of the work are as follows:

1) A comprehensive groundnut ontology has been developed, capable of answering user queries. This ontology was created from scratch, ensuring it is rich in relevant agricultural concepts.
2) An interactive interface has been created that allows users to submit queries in natural language and receive responses in natural language.

The organization of this paper is as follows: Section II reviews related work I, focusing on existing web interfaces for Indian farmers and their relevance to agricultural support. Section III provides related work II, surveying the latest developments in chatbot technology with applications in agriculture. Section IV describes the proposed model, explaining its design and how it works to meet farmers' needs. Section V details the creation of the Groundnut Ontology, covering the processes of data collection, concept identification, and structuring. Section VI presents the experiment and setup, including the RDF knowledge graph representation, Neo4j query configurations, and the development of an interface to translate user questions into queries. Section VII provides a comprehensive result discussion, evaluating the system's performance and limitations. Finally, Section VIII concludes the paper, summarizing key findings and suggesting directions for future research.

[1] https://kvk.icar.gov.in/API/Content/PPupload/k0447_28.pdf
[2] https://farmer.gov.in/
[3] https://mkisan.gov.in/
[4] http://faq.ikhedut.aau.in/1
[5] https://play.google.com/store/apps/details?id=com.aau.in.oneapp
[6] https://play.google.com/store/apps/details?id=com.aau.in.magfadi
[7] https://play.google.com/store/apps/details?id=com.aau.in.groundnut
[8] https://play.google.com/store/apps/details?id=com.khedutmall.app&hl=en_US&pli=1

II. Related Work: I Existing Web Interface for Farmers (in India)

In India, a plethora of agriculture-related applications have emerged, each designed to support farmers with a variety of essential services. These applications can be broadly categorized into simple agroadvisory systems[5] and more advanced agroadvisory systems with semantic search capabilities.

A. Categorization of App Based on Purpose

To understand these applications comprehensively, it is crucial to classify them based on the specific services they provide. The apps have been categorized according to their primary functions: weather and market price information, crop variety details, pest control, agro advisories, crop insurance information, government schemes, and agricultural news. This classification allows us to explore the unique features and benefits of each app while highlighting their relevance to different aspects of farming. Fig. 1 illustrates the categorization of applications that has been performed.

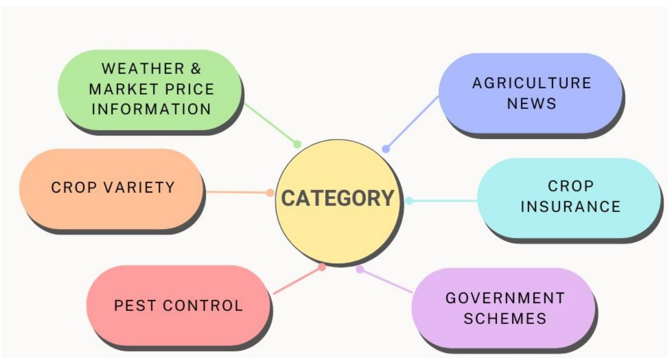

Fig. 1. Categorization of apps.

A survey of 156 agriculture-related mobile applications was conducted, and they were categorized based on their primary purpose. These applications span across various functionalities that support farmers' decision-making processes. For instance, 13 apps focus on weather forecasting, which helps farmers anticipate climate changes and plan their agricultural activities accordingly. Another 17 apps provide market information, offering crucial insights on pricing trends and helping farmers make informed sales decisions. Additionally, 16 apps leverage AI and IoT technologies, presenting innovative solutions for precision farming and resource management.

Crop insurance and government schemes are covered by 7 apps, guiding farmers on available schemes and providing them with financial protection. The largest group, consisting of 90 crop-based applications, is further subdivided into three specific areas: 25 apps offer crop information, helping farmers access detailed data on different crop varieties and innovative farming techniques; 12 apps focus on pest control and crop protection, providing strategies to identify, prevent, and treat pest infestations; and 7 apps are dedicated to soil health, offering insights on soil management practices that enhance crop yields.

This broad classification helps in understanding how digital tools can be leveraged in the agriculture sector. Weather forecasting apps, for example, assist in mitigating risks posed by unpredictable climate conditions, while market apps offer insights into the best times to buy or sell produce. Crop-specific applications provide specialized support, particularly in protecting crops from pests and ensuring soil health, which is essential for sustainable farming.

Starting with simple agroadvisory systems, these applications primarily serve as information portals, providing crucial updates on weather forecasts and market prices. Apps such as Kisan Suvidha[9] [6], IFFCO Kisan[10] [7], Agri App[11], Agri Market[12], eNAM[13] [14], mKisan[15], Ekgaon[16], myAgriGuru[17], Kisan Gujarat, AgriGujarat, and Gujarat Farm are instrumental in ensuring that farmers receive timely and relevant information about weather conditions and market trends. By providing essential data, these applications help farmers plan their activities and manage risks associated with weather and market fluctuations.

For crop variety information, the Pusa Krishi app [8] stands out. It offers insights into innovative farming techniques, crop varieties, and resource-saving technologies, which are invaluable for improving crop yields. Farmers can access detailed information about different crop varieties, helping them choose the best options for their specific conditions.

For pest control, several applications offer expert advice and practical tips to manage and mitigate pest infestations. Kisan Suvidha[18], Kheti-Badi[19], AgroStar[20], and Fasal[21] are notable examples. These apps provide specific information on pest identification, prevention strategies, and treatment options, helping farmers protect their crops from potential damage.

Crop insurance information is covered by the Crop Insurance app[22], which offers detailed information on various crop insurance schemes available to farmers. The app helps farmers understand their insurance options, eligibility criteria, and the claims process, providing financial protection against crop losses caused by unforeseen events.

Government schemes are another crucial area where mobile applications play a significant role. Apps like mKisan[23] [9], AgriMedia[24] [10], and Kisan Yojana [10] provide detailed information about various government schemes, subsidies, and benefits available to farmers. These apps ensure that farmers are well-informed about the support they can receive from the

[9] https://vikaspedia.in/agriculture/ict-applications-in-agriculture/kisan-call-center-app
[10] https://play.google.com/store/apps/details?id=com.IFFCOKisan
[11] https://apps.mgov.gov.in/details?appid=1525
[12] https://apps.mgov.gov.in/details;jsessionid=8EBEA4C94DB07B53B4FB03A49623D6CF?appid=989
[13] https://enam.gov.in/web/mobile-app
[14] https://play.google.com/store/apps/details?id=in.gov.enam
[15] https://mkisan.gov.in/Alpha/aboutmobileapps.aspx
[16] http://www.ekgaon.net/index.php
[17] https://climateasap.org/directory/myagriguru/
[18] https://kisansuvidha.gov.in/
[19] https://play.google.com/store/apps/details?id=com.freeappartist.khetiwadi&hl=en_US
[20] https://play.google.com/store/search?q=agrostar&c=apps
[21] https://play.google.com/store/search?q=Fasal&c=apps
[22] https://pmfby.gov.in/
[23] https://mkisan.gov.in/
[24] https://play.google.com/store/search?q=agrimedia+app&c=apps&hl=en-IN

government, enhancing their access to financial and technical assistance.

For agricultural news, apps such as IFFCO Kisan[25], Agri App[26], AgriMarket[27], eNAM[28], AgriBuzz[11], Kisan Yojana, Krishi Network[29], Gujarat Agri, and Gujarat Farm keep farmers updated with the latest developments in the agricultural sector. These platforms provide news on policy changes, market trends, technological advancements, and success stories, fostering an informed farming community.

The integration of semantic web technologies and ontologies is pivotal in addressing the challenges of inconsistent data and knowledge gaps in the agricultural sector.

B. Ease of Searching for Crop Product Information

An agroadvisory system allows farmers to write their questions in the system, and an agriculture expert will answer them. Farmers can also directly connect with an agriculture expert through a call to present their queries and get answers. Some existing agroadvisory applications like eSagu [12], aAQUA, and mKrishi[30] offer these features.

However, these apps have some drawbacks. A major issue is the lack of instant, personalized responses to farmers' questions. Since they do not use semantic search, farmers often have to wait for an expert to answer their specific questions, which can delay important decisions.

Semantic search represents a significant advancement in information retrieval systems [13], especially in the agricultural domain. For farmers, this means faster and more precise answers to their specific agricultural questions. Semantic search can quickly analyze and retrieve relevant data from vast databases, significantly reducing the time farmers spend waiting for answers. This immediacy is crucial for timely decision-making in farming practices.

III. Related work: II (Survey Chatbot Technology)

The history of chatbots dates back to the 1960s when Joseph Weizenbaum created ELIZA [14], the first computer program to initiate communication between humans and computers. It used a pattern-matching method to simulate human conversation. Later, in the year 1972, PARRY[31] was introduced by a psychiatrist Kenneth Colby which simulated the behavior of a paranoid schizophrenic.

By the 1990s, progress in natural language processing led to more sophisticated systems like Jabberwacky[32], which used AI to hold more natural conversations. However, the launch of Siri in 2011 was a game-changer, introducing voice-activated virtual assistants to the mainstream. Since then, AI and machine learning have propelled chatbot technology forward, leading to the development of highly advanced agents like Alexa, Google Assistant, and ChatGPT. These modern chatbots can now understand and generate human language with impressive accuracy. Chatbots play a crucial role in industries ranging from customer service to healthcare, enhancing the efficiency and smoothness of interactions with machines [15].

There has been a significant advancement in the area of Artificial Intelligence, Machine Learning and Natural Language Processing in recent years. The development in these areas have brought a marked change in various industries such as education, scientific research, medical health, including agriculture [16]. Farming is the primary source of income for millions in India. With growth in the field of AI, chatbots have emerged as innovative tools to help farmers make better decisions by providing them with access to real-time information. The technology of Chatbot applications have evolved from simple rule-based systems to advanced AI driven models [17]. This literature review highlights the development of chatbot technologies in Indian agriculture, focusing on the methodologies and innovations that have shaped the field.

A. Chatbot Technology in Indian Agriculture

Farmers can benefit from receiving correct and timely information about various aspects of agriculture, such as crop recommendations, plant disease identification, etc. A solution to this was devised by building conversational systems, which allow farmers to obtain timely answers to their queries. In 2015, AGRI-QAS [18] was developed to address farmers' queries related to crop recommendations, plant disease identification, and more. This marked the earliest advancement in this area, utilizing an index-based search technique.

In 2017, ADANS (Agriculture Domain Question Answering System) [19] introduced a significant improvement by utilizing ontology-based technology. It performed answer retrieval on a structured agriculture database, efficiently identifying relationships between agricultural concepts and providing more reliable and accurate responses to farmers' queries.

In 2018, FarmChat [20] was introduced as a conversational agent containing two user interfaces: one with Audio Only and the other with Audio+Text. It used Google's Speech-to-Text for speech conversion and used language model for query intent and entity identification, subsequently retrieving the appropriate response from the knowledge base.

AgronomoBot [21] used sensor networks to gather information about the agricultural production chain in a specific area. It integrated its information in Telegram Bot API. This marked an early integration of AI with messaging platforms to provide farmers with data-driven insights.

In 2019, AgriBot [22] provided functionalities such as crop recommendations based on current conditions, current weather details, and future weather predictions. For crop recommendations, it used algorithms such as K-Nearest Neighbors (KNN), Random Forest, and Decision Trees. The chatbot provided

[25]https://www.iffcokisan.com/agritech
[26]https://play.google.com/store/apps/details?id=com.criyagen&hl=en_IN&pli=1
[27]https://apps.mgov.gov.in/details;jsessionid=8EBEA4C94DB07B53B4FB03A49623D6CF?appid=989
[28]https://www.enam.gov.in/web/
[29]https://play.google.com/store/apps/details?id=com.krishi.krishi&hl=en_IN
[30]https://www.tatatrusts.org/our-work/livelihood/agriculture-practices/mkrishi
[31]https://en.wikipedia.org/wiki/PARRY
[32]https://en.wikipedia.org/wiki/Jabberwacky

response to user queries by accessing the Krishi Call Center database[33].

Another form of AgriBot [23] was released in 2019, utilizing a Sen2Vec-based NLP technique to answer farmers' queries. This represented a significant step forward in AI-driven agricultural chatbots, enabling more sophisticated and context-aware responses to queries.

The integration of NLP into chatbot systems has pushed the boundaries of what these technologies can achieve in agriculture. In 2020, another version of Agribot [24] incorporated an LSTM-based model to provide answers to user queries. It additionally used a CNN-based model was used to classify plant diseases based on images. The system was trained on the Krishi Call Center dataset.

By 2021, chatbots such as Krushi—The Farmer Chatbot [25] began utilizing the RASA NLU framework, an advanced NLP tool designed for processing queries in local languages. This framework identifies the intent behind each query. For weather-related queries, the system uses the appropriate OpenWeather API key to provide real-time responses. For other types of queries, it matches key entities with the database to generate suitable responses. This approach enhanced the system's ability to address farmers' needs by leveraging insights from previous interactions in the KCC datasets. It is also integrated into WhatsApp [25]. Another such similar work methodology can be seen in AgroBot where they have made use of NLP to identify the intent of user query to provide appropriate response [26].

During this period, Agroxpert addressed user queries by employing the Levenshtein distance formula. The authors compiled a dataset consisting of user queries and responses. Subsequently, user queries were matched against this dataset using the Levenshtein distance formula, allowing for appropriate responses to be generated. In instances where the system could not confidently provide an answer, the queries were escalated to human experts, creating a continuous feedback loop between AI and human expertise [27].

The usage of Artificial Neural Networks for crop disease prediction also increased during this time. Many research work were focused on providing assistance to farmers in identifying crop disease using Artificial Neural Network.

By 2023, chatbots like the Agriculture Assistant Chatbot [28] integrated a CNN-based algorithm that enabled farmers to upload crop images for disease diagnosis, offering potential remedies as well as essential information such as soil and rainfall data. Another notable work can be seen in [29], where a VGG-16-based model was incorporated for identifying diseases in plants. They provided crop recommendations based on current conditions using machine learning algorithms.

In 2024, the AI-Powered Decision Support System for Sustainable Agriculture utilized LLMs to process unstructured user queries and provide constructive farming advice. It addressed various issues, such as pest control, by using real-time data for analysis and crop management [30].

Other projects, like ChatAgri (2023) [31], explored the cross-linguistic potential of LLMs for agricultural text classification and provided end to end question answering system [32].

The development of agricultural chatbots in India has progressed rapidly over the past decade, evolving from basic AGRI-QAS to advanced AI-driven models. With the integration of LLMs such as ChatGPT and domain-specific improvements like ChatAgri [31], chatbots are becoming indispensable tools for modern farming. They offer farmers tailored, real-time solutions to the daily challenges of agriculture, positioning themselves to be vital in the future of farming.

IV. Proposed Model

Fig. 2 illustrates the workflow of the proposed question-answering system, which retrieves information from an ontology graph database.

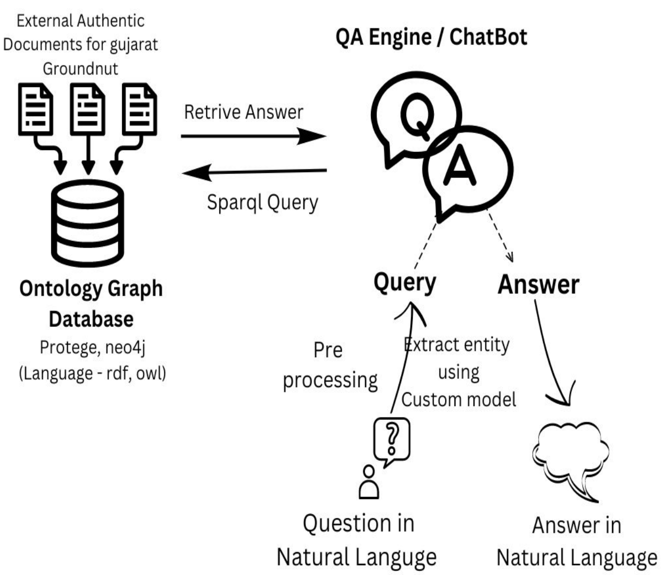

Fig. 2. Proposed work.

- Ontology Graph Database
    - Ontology Modeling: Using tools like Protégé, an ontology was created to represent the identified concepts, their properties, and the relationships between them. The ontology follows a hierarchical structure with clearly defined classes, subclasses, and individuals, ensuring that the agricultural knowledge is represented in a logical and organized manner.
    - Knowledge Graph Construction: The ontology was then translated into a knowledge graph using Neo4j, a graph database that efficiently manages the interconnected data. The knowledge graph encodes entities (e.g. "Early_Leaf_Spot X") as nodes, while the relationships (e.g. "isControlledBy Y") are represented as edges.
- QA Engine/Chatbot: The user interacts with the QA system through natural language queries. These

[33] https://www.data.gov.in/datasets_webservices/datasets/6622307

queries are processed to extract relevant entities (using a custom model) and converted into structured queries.

- Named Entity Recognition: A custom Named Entity Recognition (NER) model was developed using the pre-trained en_core_web_lg model from spaCy to meet the specific needs of the agriculture domain. While the en_core_web_lg model offers strong general purpose capabilities, it does not focus on agriculture related terms. Therefore, it was fine tuned to recognize entities relevant to agriculture, such as crops, pests, fertilizers, and diseases. Text data related to agriculture was collected, covering topics like groundnut seed variety, pest control, production technologies, and protection technologies. This data was manually labeled with agriculture-specific categories, such as Seed Variety, Controlled Pests, Diseases, and Symptoms. A simple JSON annotation methods were used for tagging. The en_core_web_lg model served as a starting point because it already contains strong word embeddings and pre-trained NER capabilities. It was fine tuned using the labeled agriculture dataset to make it suitable for identifying domain-specific terms.
- SPARQL Query: The system uses a set of predefined query templates that are dynamically adapted based on the entities and relationships identified in the user's query. By adding the identified entities (like seed variety, disease, etc.) into these templates, the system creates specific SPARQL queries to find the most relevant information from the knowledge graph.
- Answer Retrieval: The system retrieves the most relevant answers from the ontology graph database by executing the generated SPARQL query. These answers are then processed and converted into clear, natural language responses to ensure they are easily understood by the user. In cases where the query does not provide enough information to generate a precise answer, the system offers default responses.

The proposed framework integrates semantic web technologies to provide accurate, context-aware information to farmers, specifically focusing on the groundnut crop in Gujarat. By combining ontology-based knowledge representation with natural language processing, it effectively addresses the challenge of delivering precise, region-specific agricultural knowledge.

V. Creation of Groundnut Ontology

Currently, several groundnut ontologies are available, such as the AgroPortal and Agropedia groundnut ontologies. The question is whether an existing groundnut ontology can be used for gujarat-based agriculture. If so, can these agriculture ontologies be applied as they are, or will changes and modifications be needed? To address this question, detailed research was conducted by examining two existing groundnut ontologies: the agro-portal groundnut ontology[34] [33] and the agropedia [34] groundnut ontology.

As a result, it was found that the existing groundnut ontologies offer a variety of concepts that can be directly applied to the Gujarat region. Agropedia groundnut ontology is an Indian ontology. Therefore, the majority of the concepts (85%) are those that can be directly acquired from the agropedia groundnut ontology for the gujarat-based groundnut ontology[35]. In the Agroportal ontology[36], there are 700+ concepts, of which 108 concepts can be acquired for the Gujarat-based groundnut ontology. Certain concepts were found to be missing, so specific concepts related to Gujarat groundnut were added like Seed varieties (specifically used in gujarat area) Abnormality (Color(leaf), Groundnut stage, Abnor part, Shape (leaf), Symptoms), Resistance, etc.

To address the missing concepts, authentic online sources were used as references to build the ontology, including Junagadh Agriculture University[37], Anand agriculture university[38], Gujarat State seeds corporation limited[39] and Wikipedia[40] as references to build the ontology.

The ontology was manually constructed using the Protégé tool. It includes 300 classes, 21 object properties, and 8 data properties. Additionally, 104 individuals were created, which are the basic components of the ontology. In total, the ontology contains 1,569 axioms. Fig. 3 illustrates the hierarchy of classes, individuals, object properties, and data properties within the groundnut Ontology.

The groundnut crop ontology includes two major classes: Production Technology and Protection Technology. "Production Technology" class focuses on various aspects of growing groundnut crops. It has four main subclasses: Field Preparation, Nutrient Management, Water Management, and Seed and Sowing. Among these, the Seed and Sowing subclass is especially important. It includes three specific classes: Veldi, Ardhveldi, and Ubhadi. These classes contain 20+ individuals that represent different seed varieties. Each individual includes important details such as the year of release, oil content, days to maturity, pod kernel yield, etc. Fig. 4 shows how these individuals are organized in the ontology.

"Protection technology" class deals with protecting groundnut crops from various threats. A key subclass under this is Biotic Stress, which covers 45 diseases grouped into three categories: Diseases, Insect Pests, and Weeds. Another important subclass is Controlled Pest, which lists more than 35 pest control chemicals. These chemicals are linked to specific insect pests and help in managing the damage they cause. This subclass has been carefully designed to describe how pests affect crops and which chemicals are most effective in controlling them.

The groundnut ontology provides a comprehensive framework tailored for Gujarat-based agriculture, integrating concepts from existing ontologies, like agropedia and agroportal

[34]https://agroportal.lirmm.fr/ontologies/CO_337/?p=classes
[35]Agropedia http://agropedia.iitk.ac.in/
[36]Agrovoc https://agroportal.lirmm.fr/ontologies/CO_337/?p=classes
[37]Junagadh agriculture universityhttp://www.jau.in/.
[38]Anand agriculture university http://www.aau.in/.
[39]Gujarat State seeds corporation limited http://www.gurabini.com/.
[40]GroundNut Wikipedia https://en.wikipedia.org/wiki/Peanut.

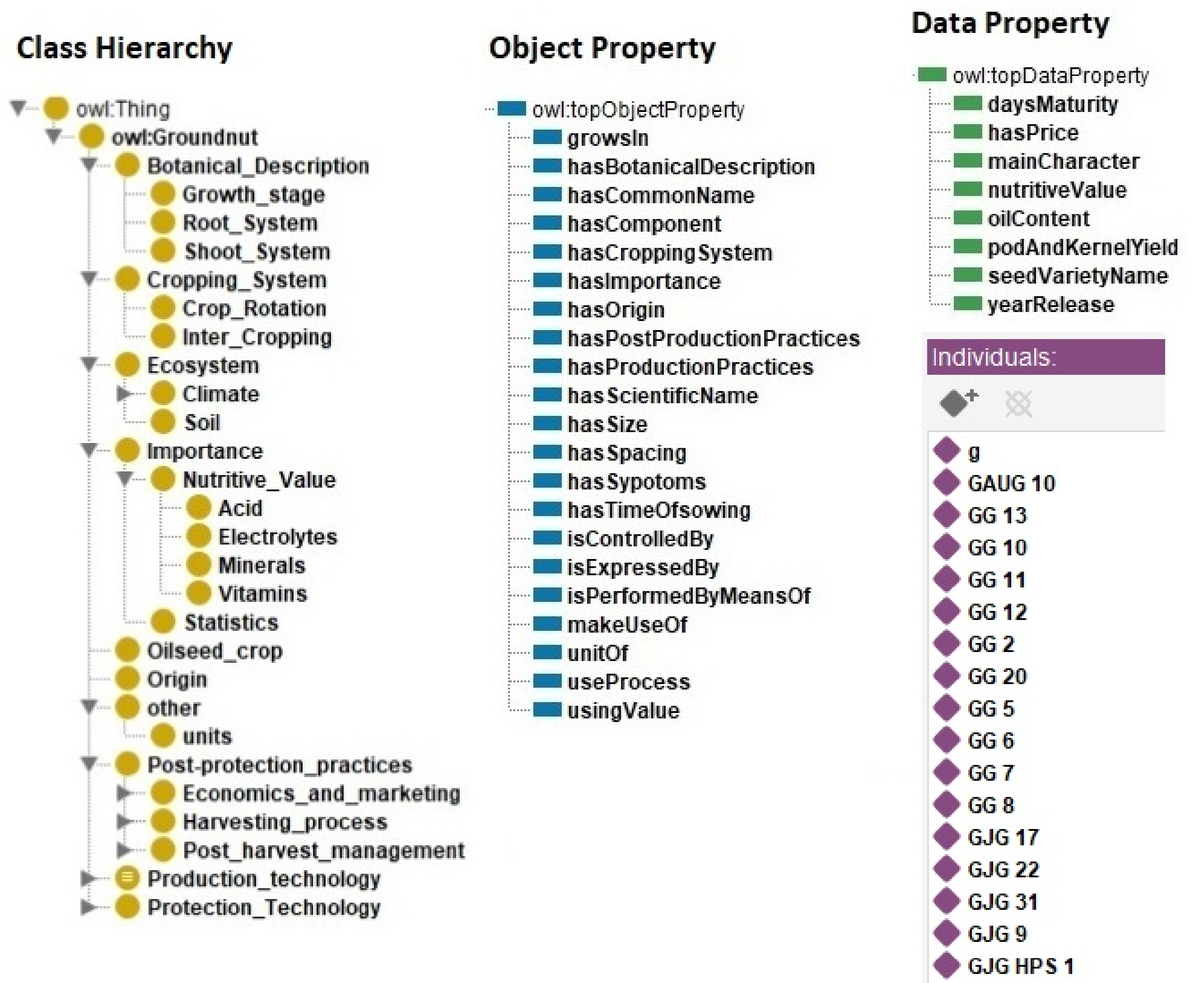

Fig. 3. Groundnut ontology class hierarchy, individuals, object and data properties.

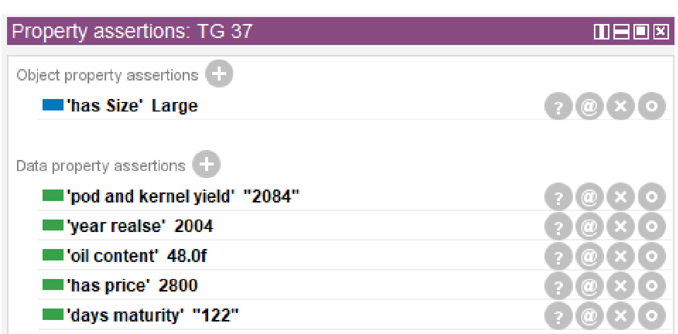

Fig. 4. TG-37 seed variety.

while addressing gaps through additional classes and properties specific to the region. Its detailed design, incorporating 300+ classes and hundreds of axioms, serves as a valuable resource for agro-advisory systems, research, and decision-making processes.

A. RDF Representation of Knowledge Graph

Neo4j[41] [35][36], the graph database [37], is used to efficiently represent and manage the groundnut ontology in the ontology-based question-answering system[42].

The groundnut ontology was first imported into the Neo4j tool, using the Neo Semantic plugin, which is required for importing ontologies into Neo4j[43]. The RDF groundnut ontology file, once imported into Neo4j, represents the fundamental components of a knowledge graph. Each entity within the groundnut ontology, such as seed variety, method of sowing, symptoms, protection technology, production technology, etc. are represented as a node in the graph. Edges connecting the respective nodes represent the relationships between these entities, which indicate dependencies, associations, and interactions. Furthermore, properties related to entities, such as daysMaturity, hasSize, hasPrice, oilContent, podAndKer-

[41] https://neo4j.com/
[42] https://neo4j.com/labs/neosemantics/
[43] https://neo4j.com/labs/neosemantics/

nelYield etc., are embedded within the graph nodes. The given query 1 demonstrates how to import the groundnut ontology into Neo4j using the n10s.rdf.import.fetch procedure. The query fetches the ontology from the provided RDF file URL and imports it into Neo4j in the RDF/XML format, with specific labels assigned to classes, object properties, and data type properties.

Listing 1: Query example

```
CALL n10s.rdf.import.fetch
("https://raw.githubusercontent.com
/PurviPatel20/with-Label/main/Groun
dnut_2.0.rdf","RDF/XML",
{
classLabel : 'Category',
objectPropertyLabel: 'Rel',
dataTypePropertyLabel: 'Prop'
});
```

The execution of query 1 successfully imports the ontology, ensuring that the defined categories, relationships, and properties are accurately integrated into the Neo4j database.

VI. Experiment and Setup

A. Neo4j Query

After successfully importing the groundnut ontology into Neo4j, the database was queried using Cypher, Neo4j's query language[44]. For example, to retrieve detailed information about a specific seed variety, a sample Cypher query to extract infomation about specific seed variety is shown below in query 2.

Listing 2: Query to retrive information about seed variety

```
OPTIONAL MATCH (i:ns0__Ubhadi)
WHERE i:ns0__Ubhadi
RETURN
i.rdfs__label As Seed_variety,
i.ns0__yearRelease As year,
i.ns0__daysMaturity As Days,
i.ns0__hasPrice As Price,
i.ns0__oilContent As Oil,
i.ns0__podAndKernelYield As
pod_and_kernel
```

This query language allows for precise data retrieval from a database. The Cypher query uses an OPTIONAL MATCH clause to retrieve data about the "Ubhadi" seed variety from a database. The WHERE clause ensures that only nodes labeled "ns0_Ubhadi" are returned in the query results.The RETURN statement indicates which properties will be included in the output, as illustrated in the Fig. 5. These properties encompass the seed variety label, release year, days to maturity, market price, oil content, and pod and kernel yield. This structured approach enables the extraction of comprehensive information about the "Ubhadi" seed variety.

[44]https://neo4j.com/docs/cypher-manual/3.5/

Seed_variety	year	Days	Price	Oil	pod_and_kernel
"GJG 9"	2009	"103"	3410	49.0	"1663"
"TPG 41"	2004	"122"	null	49.0	"2088"
"GG 8"	2006	"104-107"	null	46.0	"1776"
"GG 7"	2001	"100"	null	49	null
"GG 6"	1999	"115-120"	null	50.3	"2.78"
"GG 5"	1996	101	2310	49.2	"1270"
"TG 37A"	2004	"122"	2800	null	null
"GG 2"	null	"100-105"	2310	null	null
"J 11"	null	"110-115"	null	null	null
"GG 20"	1992	"120"	2310	50.7	"1960"

Fig. 5. Neo4j query to get all details of groundnut seed variety.

B. Interactive Interface for Question to Query Conversion

Neo4j was used in conjunction with Google colab, which allowed the use of important libraries such as py2neo, neo4j-driver, spaCy, and Gradio. The goal was to create a chatbot-style interface for the groundnut ontology. This given algorithm receives a user's question as a natural language (string input) and outputs a response string.

Algorithm 1 Question Processing and Answer Extraction

1: **Preprocessing:**
- a. Tokenize the question using the NLTK library's `word_tokenize` function.
- b. Perform part-of-speech tagging on the tokenized words using the NLTK library's `pos_tag` function.
- c. Extract the entity from the question using the `extract_entity` function (custom model).

2: **Answer Extraction:**
- a. If the extracted entity is None, return a default message indicating that the seed variety is not understood.
- b. Open a session with the Neo4j database using the `driver.session()` function.
- c. Identify entities from the question.
- d. Execute a Neo4j Cypher query to retrieve an appropriate response.
- e. If the query returns no results, return a default message indicating that no information is found.

The input text may contain important entities such as seed variety, disease name, symptoms etc. To identify these important named entities, the en_core_web_lg pipeline from spaCy was used.A custom model was trained to recognize entities such as “Price”, “Rate”, “Disease”, “Symptoms”, “Cure”, “Pesticide”, “Crop”, etc. This model was designed to handle specific cases that were not covered by the standard model.

The standard model, when given with sentence like “What is a price of GJG 22?” identified “GJG 22” as an organization. Similarly, the model was not able to identify name of the symptoms. It classified it as a product entity. To address this, a custom dataset was created to identify the text “GJG 22” as a crop entity. Using Gradio, a python library, a system was built that allows users to ask questions in natural language and receive answers based on the groundnut ontology. This approach made it easier for users to interact with the Neo4j database and get useful information.

VII. Results and Discussion

The provided Fig. 6 depicts a sample example of an ontology-based question-answer system. In the example presented, the user input is a question formulated in natural language: “What is the price of GJG 22?”. The system processes this query, which utilizes ontological knowledge to understand and generate an appropriate response. The response generated by the system is displayed within the image: “The price of GJG 22 is 2310.”

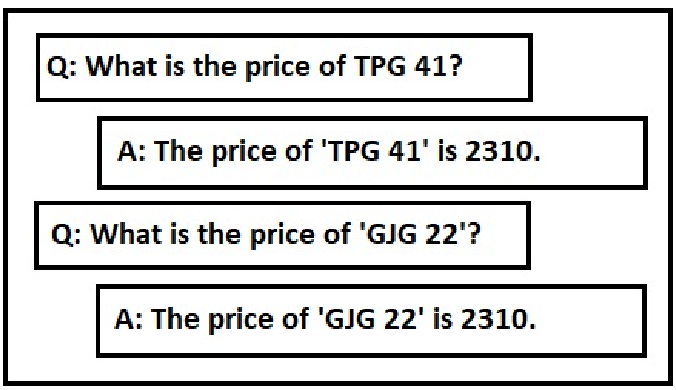

Fig. 6. Output 1: Snapshot of an ontology-based question-answer.

In Fig. 7, the user inputs a query in natural language, specifically requesting comprehensive details about the seed variety labelled “GJG 22.” The system then generates an appropriate response, which is displayed in the form of text within the interface. The response provided by the system encapsulates various attributes associated with the seed variety “GJG 22.” These attributes include: price, days maturity, pod and kernel yield and oil content. This interaction demonstrates the system’s ability to interpret complex natural language queries and retrieve structured information from groundnut ontology, facilitating efficient access to relevant data for users.

In the evaluation of the Question Answering (QA) system, a total of 100 distinct questions were submitted, and the responses were manually ranked based on their correctness. In this ranking scheme, a rank of 5 indicates a fully correct answer, while a rank of 1 represents a fully incorrect response. Ranks 2 to 4 denote varying degrees of partial correctness, reflecting the extent to which the answers met the query requirements. The results of the evaluation are summarized in the Table I.

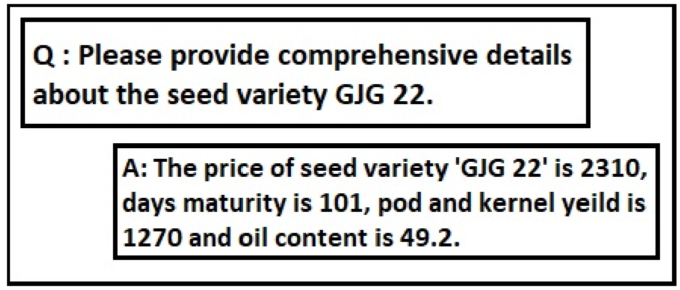

Fig. 7. Output 2: Snapshot of an ontology-based question-answer.

TABLE I. Ranking Table

Rank	No. of Questions
5	50
4	20
3	20
2	5
1	5

For instance, in the case of rank 1, the system failed to recognize “Peanut Strips” as a single entity, which significantly hindered its ability to retrieve the appropriate response from the underlying groundnut ontology. Additionally, an example of a partially correct response can be observed in the question, “What are the pesticide methods for the following symptoms: Buckling and crinkling between veins?” In this instance, the system provided an answer that was relevant but lacked completeness.

$$\text{Average Rank} = \frac{\sum_{i=1}^{n}(r_i \times q_i)}{\sum_{i=1}^{n} q_i} \tag{1}$$

Where:

- n is the total number of different ranks (which is 5 in this case).
- $r_i \times q_i$ represents the weighted contribution of each rank to the total score.

Using the formula 1 for average rank, the overall performance of the system was calculated to be 4.05. This indicates that while the system provided a significant number of accurate responses, there is still considerable room for improvement in handling complex queries and recognizing key entities.

VIII. Conclusion and Future Work

The proposed ontology-based QA system can be a valuable resource for farmers in Gujarat. Farmers can ask questions in natural language, and the system is designed to provide relevant answers using the groundnut ontology. The main objective is to give farmers access to insights that can help improve their farming practices and enhance groundnut crop yield. The overall accuracy of the answers is 80%. Additionally, the ontology aids in semantic disambiguation. In the future, more concepts can be added to the ontology. The proposed model can also be adapted for other crop ontologies and developed in different languages.

References

[1] S. Chaudhary, M. Bhise, A. Banerjee, A. Goyal, and C. Moradiya, "Agro advisory system for cotton crop," in *2015 7th International Conference on Communication Systems and Networks (COMSNETS)*, 2015, pp. 1–6.

[2] M. Fernandez, V. Lopez, M. Sabou, V. Uren, D. Vallet, E. Motta, and P. Castells, "Semantic search meets the web," in *2008 IEEE International Conference on Semantic Computing*, 2008, pp. 253–260.

[3] D. Martyniuk, N. Karam, M. Falkenthal, Y. Dong, and A. Paschke, *Semantifying the PlanQK Platform and Ecosystem for Quantum Applications*, 09 2023.

[4] A. Daud, M. H. Ullah, A. R. Banjar, and A. A. Alshdadi, "Ontological modeling and semantic search in quran," *IJCSNS*, vol. 22, no. 5, p. 771, 2022.

[5] A. Ga, A. Mukherjee, M. Roy, and N. Chandra, "Mobile based agro advisory service and farmer's willingness to pay: A case study in bageshwar district of uttarakhand," vol. 16, pp. 976–986, 12 2021.

[6] M. J. Bisheko and R. G, "A study on farmers' perceptions about the scope of the kisan suvidha app in improving agricultural sustainability," in *2023 Conference on Information Communications Technology and Society (ICTAS)*, 2023, pp. 1–5.

[7] *"Case Study IFFCO Kisan Agriculture App Evolution to Data Driven Services in Agriculture_2016"*. GSMA, Oct 2016.

[8] A. Yadav, P. Thilagam, and S. Singh, "Mobile applications for agricultural transformation: Types, impacts, case studies, and recommendations," 2023.

[9] *"Case Study mKisan, India"*. GSMA, Feb 2015.

[10] S. K. Mahapatra and J. Pradhan, "Smartphone apps-an eminent ict tools for agricultural development & agri-information dissemination," *ICT: A Catalyst to Transform Rural India*, p. 5.

[11] D. B. J. Rabba Gundu Devender Goude, "Agribuzz-agriculture management system," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, p. 5, Apr 2022.

[12] P. K. Reddy, G. V. Ramaraju, and G. S. Reddy, "esagu™: a data warehouse enabled personalized agricultural advisory system," in *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 910–914. [Online]. Available: https://doi.org/10.1145/1247480.1247586

[13] V. UREN, Y. LEI, V. LOPEZ, H. LIU, E. MOTTA, and M. GIORDANINO, "The usability of semantic search tools: a review," *The Knowledge Engineering Review*, vol. 22, no. 4, p. 361–377, 2007.

[14] J. Weizenbaum, "Eliza—a computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 9, no. 1, p. 36–45, Jan. 1966. [Online]. Available: https://doi.org/10.1145/365153.365168

[15] P. Y. Niranjan, V. S. Rajpurohit, and R. Malgi, "A survey on chat-bot system for agriculture domain," in *2019 1st International Conference on Advances in Information Technology (ICAIT)*, 2019, pp. 99–103.

[16] W. Maroengsit, T. Piyakulpinyo, K. Phonyiam, S. Pongnumkul, P. Chaovalit, and T. Theeramunkong, "A survey on evaluation methods for chatbots," 03 2019, pp. 111–119.

[17] S. Singh and H. K. Thakur, "Survey of various ai chatbots based on technology used," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2020, pp. 1074–1079.

[18] S. Gaikwad, R. Asodekar, S. Gadia, and V. Attar, "Agri-qas question-answering system for agriculture domain," 08 2015, pp. 1474–1478.

[19] M. Devi and M. Dua, "Adans: An agriculture domain question answering system using ontologies," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017, pp. 122–127.

[20] N. Darapaneni, R. Tiwari, A. R. Paduri, S. Saurav, R. Chaoji, and Sohil, "Farmer-bot: An interactive bot for farmers," *ArXiv*, vol. abs/2204.07032, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:248178166

[21] G. Mostaco, L. Campos, I. Souza, and C. Cugnasca, "Agronomobot: a smart answering chatbot applied to agricultural sensor networks," 06 2018.

[22] D. Sawant, A. Jaiswal, J. Singh, and P. Shah, "Agribot - an intelligent interactive interface to assist farmers in agricultural activities," in *2019 IEEE Bombay Section Signature Conference (IBSSC)*, 2019, pp. 1–6.

[23] N. Jain, P. Jain, P. Kayal, J. Sahit, S. Pachpande, J. Choudhari, and M. Singh, "Agribot: Agriculture-specific question answer system," 06 2019.

[24] B. Arora, D. S. Chaudhary, M. Satsangi, M. Yadav, L. Singh, and P. S. Sudhish, "Agribot: A natural language generative neural networks engine for agricultural applications," in *2020 International Conference on Contemporary Computing and Applications (IC3A)*, 2020, pp. 28–33.

[25] M. Momaya, A. Khanna, J. Sadavarte, and M. Sankhe, "Krushi – the farmer chatbot," 06 2021, pp. 1–6.

[26] C. Bhuvaneswari, H. Pokhariya, P. Yarde, V. Vekariya, H. Patil, and N. L, "Implementing ai-powered chatbots in agriculture for optimization and efficiency," 01 2024, pp. 1–7.

[27] "Agroxpert - farmer assistant," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 506–512, 2021, international Conference on Computing System and its Applications (ICCSA- 2021). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666285X21000443

[28] K, E. Venkat Reddy, K. Sathvik, S. Laya, and Harsha, "Agriculture assistant chatbot," *International Journal of Innovative Science and Research Technology (IJISRT)*, vol. 5, pp. 116–123, 2024.

[29] A. Marla, R. Paul, A. K. Saha, N. K. Basha, and B. Anandhakrishnan, "An agrobot: Natural language processing based chatbot for farmers," in *2023 4th International Conference on Smart Electronics and Communication (ICOSEC)*, 2023, pp. 1235–1241.

[30] E. Asolo, I. Gil-Ozoudeh, and C. Ejimuda, "Ai-powered decision support systems for sustainable agriculture using ai-chatbot solution," *Journal of Digital Food, Energy and Water Systems*, vol. 5, 06 2024.

[31] B. Zhao, W. Jin, J. Del Ser, and G. Yang, "Chatagri: Exploring potentials of chatgpt on cross-linguistic agricultural text classification," *Neurocomputing*, vol. 557, p. 126708, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231223008317

[32] N. Chandolikar, C. Dale, T. Koli, M. Singh, and T. Narkhede, "Agriculture assistant chatbot using artificial neural network," in *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, 2022, pp. 1–5.

[33] C. Jonquet, A. Toulet, E. Arnaud, E. D. Yeumo, J. Graybeal, M.-A. Laporte, M. A. Musen, P. Larmande, S. Aubin, V. Pesce, and V. Emonet, "Agroportal: A vocabulary and ontology repository for agronomy," *Computers and electronics in agriculture.*, vol. 144, p. 126—143, January 2018. [Online]. Available: https://doi.org/10.1016/j.compag.2017.10.012

[34] M. Sini, V. Yadav, J. B. Singh, V. Awasthi, and P. Tv, "Knowledge models in agropedia indica," 2009. [Online]. Available: https://api.semanticscholar.org/CorpusID:55110203

[35] F. Holzschuher and R. Peinl, "Performance of graph query languages: Comparison of cypher, gremlin and native access in neo4j," 03 2013.

[36] F. Gong, Y. Ma, W. Gong, X. Li, C. Li, and X. Yuan, "Neo4j graph database realizes efficient storage performance of oilfield ontology," *PLOS ONE*, vol. 13, no. 11, pp. 1–16, 11 2018. [Online]. Available: https://doi.org/10.1371/journal.pone.0207595

[37] Y. Chen and X. Xing, "Constructing dynamic knowledge graph based on ontology modeling and neo4j graph database," in *2022 5th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2022, pp. 522–525.