SPARQL Query Answering over Cotton Crop Pest Ontology

Mahesh D. Titiya, Department of Computer Engineering, Government Engineering College, Gujarat, India
Vipul A. Shah, Department of Instrumentation & Control, Dharmsinh Desai University, Gujarat, India

ABSTRACT

There is a very large amount of data related to agricultural practices being collected via different sources but it is not being used for maximum benefit by the farmers due to lack of mediums for the information to flow and other factors like language differences, lack of technology to access that information etc. Information Communication Technology (ICT) can help to bridge that gap by creating systems that are easier to access and are able to answer basic queries for the farmers that might help them in improving their farming practices that benefits everyone. Such a system should make use of all the data sources available and provide processed information that makes sense to the user. An Agro advisory system is one where farmers and other people associated with agriculture can get automatic recommendations for any and all of their queries. Considering the amount of research and development being done in the field the problem is to distribute that newly found information and research and present it to the end users so that they can leverage the power of ICT to improve their production and ways. We envision a platform where the farmers or other interested parties can get access to that information using a very simple easy to use interface.

KEYWORDS

Ontology, Semantic Web, RESTful services, Agro-advisory system, Service Oriented Architecture, Mobile application

1. INTRODUCTION

Due to the utilization of Semantic web technologies the user able to retrieve relevant information very easily. The semantic web is an extension of the current web in which information is given well-defined meaning, better enabling computers, and people to work in cooperation [19]. The semantic web gives importance to web page contents which allow computers or device to recognize contents of information in web pages as an individual can understand. It responds to user request based on the meaning of the query. The purpose of the Semantic web is to convert unstructured or semi-structured into structural data for the web[22]. Ontology organizes the information in a semantic web with a structured framework. It can be used for representing knowledge of a particular domain.

The ontology is shared dictionary which depicts the individual, domain concepts, properties of individual, and relationship between concepts. The agricultural knowledge base can be created with the help of ontology.

Unfortunately, the knowledge base for the plant and crop production is not updated regularly as well as it is not correctly utilized by farmers. Currently, the research efforts are made for developing ontology with two methods which are automatic and semi-automatic.

The cotton crop pest ontology which developed by us is going to be used by farmers to submit their query related to cotton farming. The ontology able to answer any difficult query generated by farmers and it provides a relevant solution to farmers.

2. MOTIVATION FOR RESEARCH

In the domain of agriculture, the farmers have queries regarding varieties of the crop, soil information, favorable climate condition for the crop, cultivation techniques, disease and pest affecting crop and prevention of it. Farmers express their queries in natural language which is answered by agricultural. Because of lack of time, and access it is not possible for experts to

present physically to respond each query for every farmer. Due to an absence of the expert, the farmers may not able to understand the expert's opinions or suggestions. In such scenario, there is a need for an agro advisory system which can bridge a gap between agricultural experts and farmers. We are motivated to develop an agro advisory system which can able to answer the queries of farmers which will help farmers in improving farming practices and yield production.

3. LITREATURE SUREVEY

During literature survey, we found existing Agro Advisory systems like Kissan Kerala [7], aAQUA [12], mKrishi [5].

Karshaka Information Systems Services and Networking (KISSAN) is an integrated, multi-modal delivery of agricultural information system, which provides dynamic and useful advisory services to the farming community across Kerala. It is one of the leading citizen-centric e-governance projects of Agriculture Department, Government of Kerala which is conceived, developed and managed by the Indian Institute of Information Technology. The key feature of KISSAN uses integrated service delivery model which makes expert's opinion available to farmers anywhere in the state.

The discussion forum aQUA which is designed by IIT Mumbai. The user can use the web browser to access the system. The primary goal is to create a platform for farmers to express their agricultural related problem to experts. It is multimedia and multilingual querying forum. The user can submit their agricultural queries after registration and expert will respond to farmers query.

TCS group has developed a mKrishi system which is a mobile based advisory system which provides audio-video facilities to post farmer's queries. The farmers can send the status of a crop by taking pictures of the crop using a camera of mobile phones. The farmers can also record their queries and submit to the system. mKrishi system collects climate information from nearby weather stations and soil and crop data from sensors. The system advice to farmers on their mobile phone after analyzing collected information.

4. RESEARCH CHALLENGE

We have found several research challenges to develop our system.

User Interface: To build the user interface for farmers to express their query is challenging issue. The system should be more easy to use to farmers.

Development of Advisory system based on Ontology: System should consist of knowledge base data. The knowledge base of the system should support semantic searches. Ontology is knowledge representation techniques which support semantic technologies. Such a knowledge base facilitates us to present improved responses derived from inference and reasoning ability on ontology. The main goal is to construct cotton crop ontology which can answers queries to each cotton crop farmers.

Ontology Reasoning: It is challenging issue to answer the query for user for which the information is not available in knowledge base statically. All the information in the knowledge base is not explicit. The system should have the capability to infer the knowledge from existing knowledge base. Ontology reasoner should generate a recommendation to farmers after reasoning process.

Natural language queries: The farmers express their query in natural language. A natural language query contains normal terms which used in local language by users. It does not have any specific syntax or format to express their query. The system should allow the user to enter their query in any form, including a statement, a question, or a simple list of keywords. The response of the system should be independent of the mechanism which user has preferred to submit their query.

Deployment of the system: System should be easy to access. The system should be deployed on a cloud which can be effortlessly available to users. Knowledge base, services to retrieve agricultural information and records of agricultural should be placed on cloud such that the user can able to use from anywhere, independent of location.

5. KISANMITRA: ONTOLOGY BASED RECOMMENDED SYSTEM

After reviewing existing systems and looking at the specific open research issues, we have developed an Ontology-Based Agro Advisory System for Cotton Crop Farmers named KisanMitra.

It provides flexibility to farmers to express their queries in natural language It can respond to farmer's queries such as pest, diseases, pest preventions techniques, varieties, timelines of various activities such as irrigation, fertilizing, harvesting, and insecticide spraying for cotton crop.

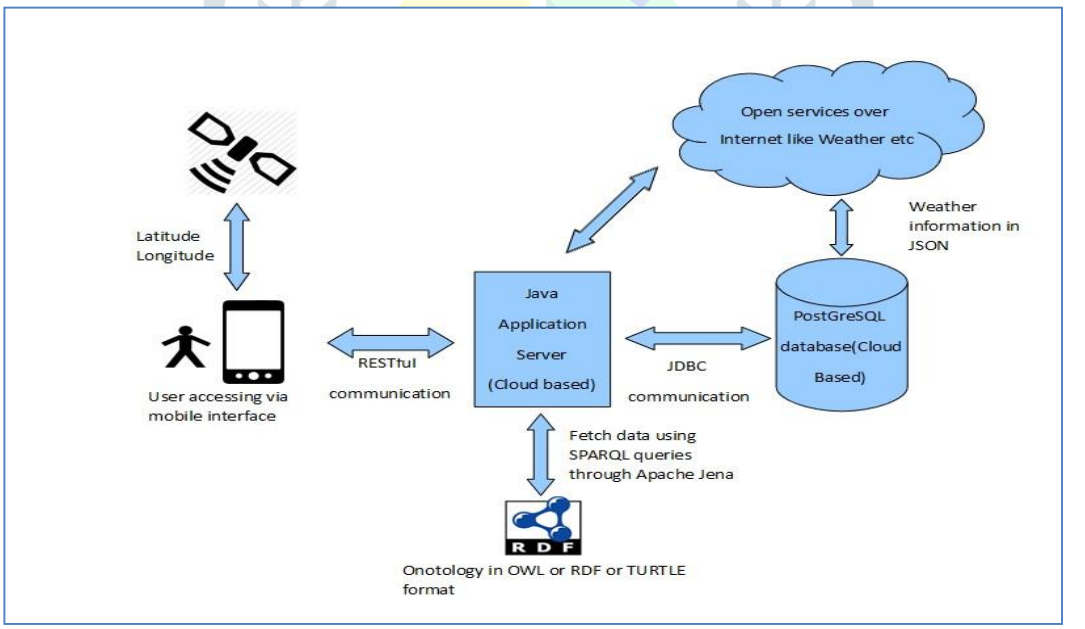
Our system basically encourages keyword-based query where a user enters a query in simple keywords which is interpreted and analyzed by the system, performs semantic searches and advice to farmer based on its context and severity. The farmers can access KisanMitra using a mobile phone or a web browser. We have developed an Android OS based mobile application. The android application is created using Java, Android SDK-version 1.4 and Eclipse IDE.

6. DESIGN AND IMPLEMENTATION OF KISANMITRA

KisanMitra is recommended system developed for cotton farmers which will help to improve cotton production. The information for the system is collected using various data sources such as Geographical Information system data, SQL data, and Resource description framework knowledge base. The used can access KisanMitra cloud-based online system using mobile device or web browser. In this system, RESTful services are developed to make communication between the system components. Figure 1 shows the architecture of KisanMitra system.

The system provides information to farmers regarding disease and pest in surrounding farms, soil health card information, climatic condition, pest, and disease etc. The system recommends to farmers for prevention and cure techniques to be used for avoidance of pest and disease to the cotton crop. The system can be accessed using user-friendly interface which requires minimal training to use the system.

Figure 1. KisanMitra System Architecture



The major components of the system are explained below.

Structured Query Language (SQL) database: The database contains the information which does not alter regularly. It contains information related to farm survey and soil. The information which contains in SQL database used for handling registration of the farmers and it helps farmers

to get relevant information. Apart from the static data, weather and soil health card data also stored in the SQL database. The SQL database stored as a PostgreSQL database on cloud-based server.

Query Engine: It is one of the core components of the system which has the capability to handle query which is raised by the farmer. To answer the query the query engine will use the knowledge base of cotton farming practices which is in the form of ontology and database. Currently, we are handling the ontology-based query. This is taken care by designing an algorithm which returns a path that best matches cluster of resources selected by a user query.

User Interface: The user can access the system using a mobile phone application or web browser. The Android application is created using Java, Android SDK - v14 and Eclipse IDE. We have designed a user interface to be very simple and easy to use without a lot of manual inputs by keeping drop-downs for selection wherever possible instead of the text fields. The farmers can submit a query to the system in natural language which is parsed, tokenized and mapped to resources in the ontology and response of query given to the farmer. For the development of the proposed system, we have followed the below sequence for tasks during development. We have followed the below sequence for tasks during development.

- Designing of databases.
- Construction of cotton crop ontology.
- Development of web services.
- Development of Android mobile application and web based interface.
- Deployment of the database and web services on cloud based provider making it accessible from the internet.
- Notification and alert generation to farmers.

7. DETAILS OF IMPLEMENTATION

We have used the following Domain specific knowledge materials and tools for development, testing and deployment:

7.1 Software and Hardware Requirements

We have used the below tools and softwares for development, testing and deployment:

- Java Development kit 1.6 or higher.
- Eclipse J2EE Juno version for developing REST based web services.
- Apache Tomcat (preferred) or other Java Application web server.
- PostgreSQL 9.2, pgAdmin, PostGIS 2.0.
- Android SDK v4.0 or higher.
- Apache Jena for RDF queries.

7.2. Cotton Crop specific knowledge material

To collect information regarding cotton farming practices we have referred various resources such as handbook of cotton [6], agropedia [2], and web resources for Central Institute for Cotton Research (CICR)[9,21], Indian Agricultural Statistics Research Institute (IASRI)[13],etc.

7.3. Agropedia

Agropedia contains knowledge models for crops. Knowledge models are useful in mapping the knowledge and tagging content of the agropedia portal so information is becoming semantic. It also provides some guidelines to create knowledge model. Knowledge models are mainly used to organize and search agriculture information. We can access the various crops knowledge models like Chickpea, Litchi, Rice, Sorghum, Sugarcane, Wheat, Safflower, Cotton, Pigeon Pea, and Groundnut from agropedia. Knowledge models represent the knowledge in terms of structure. With the help of agriculture experts, IIT Kanpur has organized the agropedia in a top-level generic map for crops which is called the “Foundational Agriculture Crop Ontology”.

7.4. RESTful Web Services Implementation

We have developed RESTful web services to establish communication between different components of a system. Some of the web services are explained below.

WeatherInfoService: This service returns current weather information of particular location for which latitude and longitude provided as an input. This service records current weather conditions of the defined location with latitude and longitude. The service is self-directed which doesn't require any human dealing. It can be used in combination of location service to get the weather conditions of the user's current location.

SoilInfoService: This service returns information about soil health such as information about presence of nitrogen, phosphorus, potassium, PH value present in soil. It also returns type of soil whether it is loamy, sandy etc

FarmInfoService: This service returns information about survey number of farm, location detail of farm, longitude, and longitude of farm based on parameters passed farm survey number and/or farmid.

DistrictInfoService: This service returns names of all districts where the major cultivated crop is cotton. It retrieved information from our database based on passed parameter districtId.

TalukaInfoService: This service returns information about taluka of farmers from our database based on passed parameters like farmerid.

VillageInfoService: This service returns information about the village of farmers based on parameters passed farmerid.

8. RESULTS

We have placed our system on cloud-based server Heroku which allows KisanMitra system to access from anywhere by using a mobile device which has internet access. We have deployed SQL database, as a PostGreSQL server on Heroku server. The ontology .owl file in the form of RDF data along with the web services file is stored on the Heroku server.

The farmers will be capable of finding the solution of a query on their mobile phones without being worried about the resource for data. The cell phone just needs to communicate with RESTful web services. The interface for the mobile phone is kept very simple and user-friendly. It has drop down menus for providing inputs for a query. The farmers do not require much training to use our system.

The cotton crop ontology is updated at regular interval by agro expert after making a login to the system. In our system, there is a provision for reporting new disease or pest observed by farmer which affecting the cotton crop. These diseases or pest-related information will be updated instantly in our ontology.

Our KisanMitra system requires information for farmer location soil, type of soil etc.To facilitate that we have prepared the basic form templates using formhub. Those templates can be used to collect data using a mobile device or web browser in offline mode as well.

9. SYSTEM SCREENSHOTS

The system is using crop pest ontology as a knowledge base. The concepts in ontology with its name are represented by a rectangle.

The cotton ontology stores various types of information such as varieties of a cotton crop, different types of soil such as loamy, sandy loamy soil ,favourable climatic conditions, pest and disease affecting cotton crop, prevention techniques to be used for avoiding disease and pest on cotton crop.

The cotton ontology contains cotton varieties information. Desi Hybrids, G.hirsutum, Intra-hirsutum Hybrids, G.arboreum, G.harbaceum, and Hirsutum Barbadense are the Cotton Varieties. Here in figure rectangle with diamond states the individuals of concepts. For example, G.Cot.DH.7 and G.Cot.DH.9 are the individuals of Desi Hybrids concept.

The climatic condition for cotton crop is also stored in cotton crop ontology. Individual of ontology represented with rectangle with diamond shape. The different concepts are created for Mean_temp,Rainfall_during_growing_season,DrySpell_Length_at_Bolldevelopment,Rainfall_during_BollDevelopment,Length_of_Growing_Period,Total_Rainfall,Min_Temp,Max_Temp,Min_Rainfall and Max_Rainfall.

There are different types of disease such as bacterial, fungal, viral affects cotton crop. The disease ontology constructed which stores the information about different types of disease, symptoms of a disease, and pesticide and insecticide which need to be used to cure and prevention of disease. The disease ontology is shown in Figure 2 and Figure 3. The concepts in disease ontology are connected to other concepts by using more than one relational name. e.g. is_prevent_by is relation name connects disease and prevention concept. Here disease is domain concept and prevention is range concept.

Figure 2. Snapshot of Disease Ontology Concepts

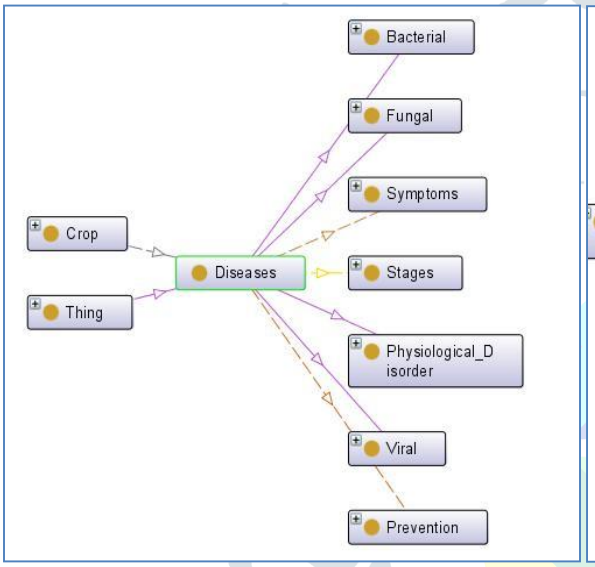
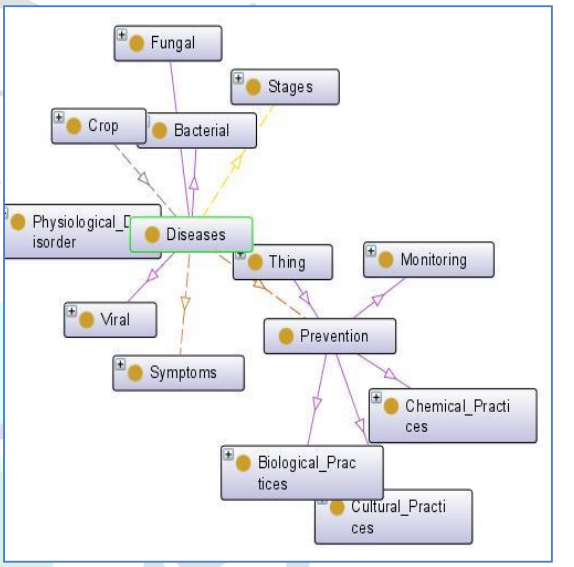


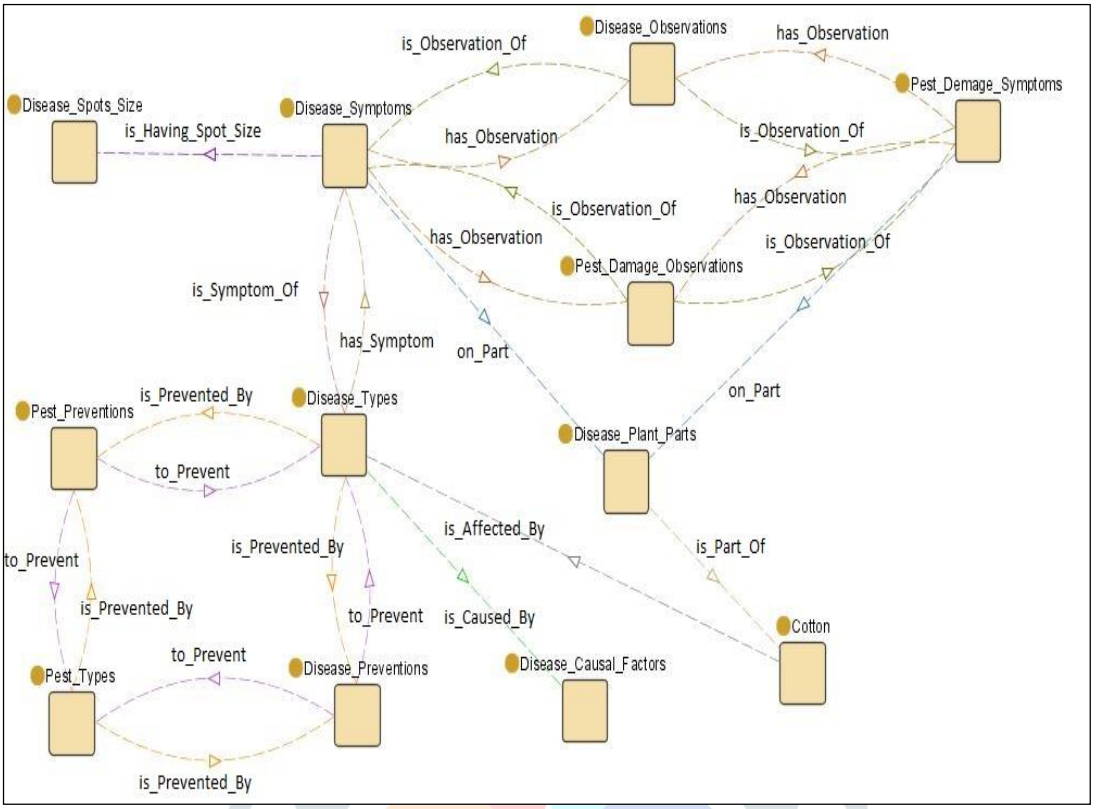
Figure 3. Snapshot of Disease Ontology Concepts



Semantic Technologies features support storing, loading and operations on RDF/OWL models. Each model contains a set of subject-object relationship triples organized as an RDF/OWL graph of directed labeled edges.

The edge is the link (or relationship) that connects a subject node to an object node and is labeled by a predicate [21]. Figure 4 shows the relationships between the two concepts. Concepts are related to other concepts by one or more relations. Here in figure link originate from the domain concept and converge into range concept. For example, for “to_Prevent” relation relates Pest_Preventions and Disease_Types. In this relation Pest_Preventions is the domain concept and Disease_Types is range concept. There is also reverse relation is_Prevented_By between these two concepts.

Figure 4. Relation between concepts



For direct querying of ontologies, SPARQL can be used. SPARQL is the query language of the Semantic Web. It queries RDF graphs, which consist of various triples expressing binary relations between resources, by specifying a sub graph with certain resources replaced by variables. SPARQL queries were written to fetch the data from knowledge base which is in the form of ontology to respond query of farmers. Few of SPARQL queries are explained below.

Query-1) Advice for disease prevention technique based on observed disease and parts of cotton plant affected by a disease.

The query retrieves the data from the ontology which is act as an input and the recommendation for disease prevention is the output generated. The recommendation is generated based on the symptoms and the affected parts of a cotton crop. Recommendation consists of pesticide, insecticide, fungicide as a disease prevention practices. The response generated by a system for query1 is shown in Figure5.

Query 2) Advice for disease prevention technique to be used for cotton plant based on current climate condition.

It will retrieve current climate data and recommends disease prevention techniques such as chemical and biological control to prevent the disease. Here the weather data taken as an input and prevention technique is an output. The response generated by a system for query2 is shown in Figure 6.

Query 3) Advice for the pesticide details to be used for Bollworm pests for cotton plant.

To get correct pesticide name for prevention of Bollworm pests we have used ontology and reasoner which inference the ontology to generate the desired result. It will generate all the results of pesticide name for bollworm pest. The response generated by a system for query3 is shown in Figure 7.

Figure 5. SPARQL Query result for Query 1

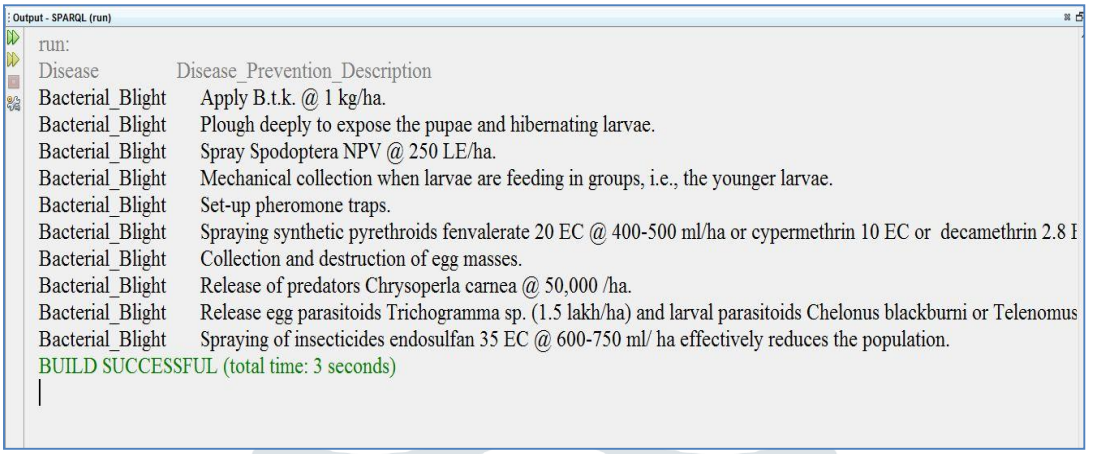


Figure. 6. SPARQL Query result for Query 2

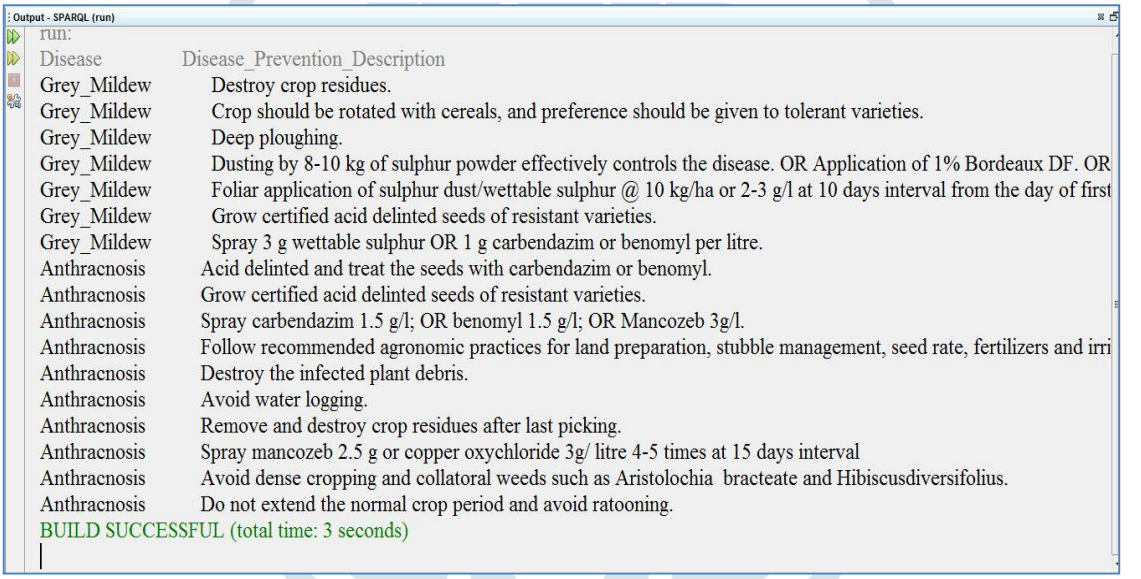
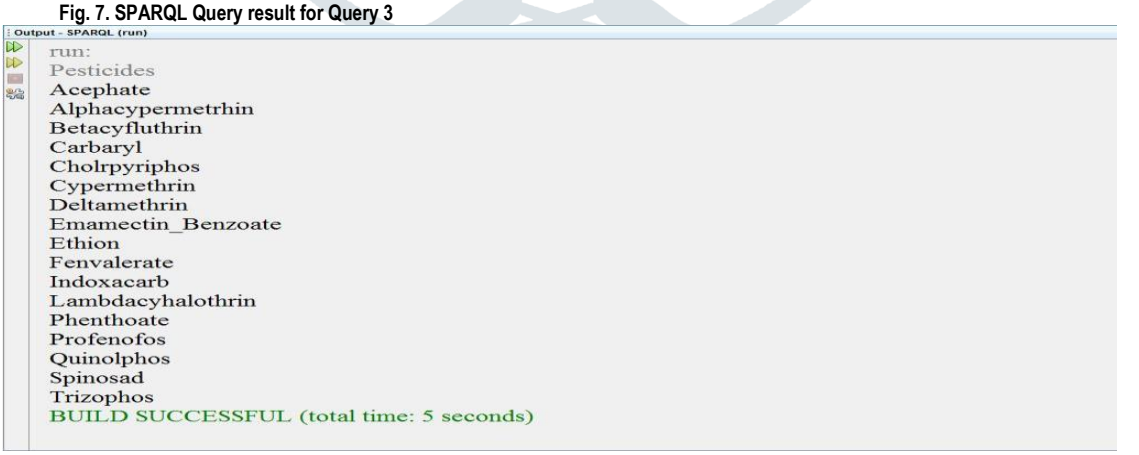


Fig. 7. SPARQL Query result for Query 3



10. CONCLUSION AND FUTURE SCOPE

Our system can be more useful to farmers if we could include current climate condition data and prediction techniques for weather. The system can generate and send alert and notification to farmers so that precautionary steps can be taken in advance before heavy damage occurs to a farmer. The system can also be extending to include the soil health condition of a farm. After analyzing soil health data the system will generate recommendation such as which is the best crop to cultivate, what is the best fertilizer to use to increase the productivity? To provide the alert for the weather change and recommendation to improve the soil condition more concepts of soil and weather should be included in the ontology which make the scope of ontology large. The system can generate effective recommendation if it has a larger ontology and efficient inference engine capability.

Cotton crop ontology is a very useful resource for processing Cotton agricultural knowledge base. Our KisanMitra system is pioneer and pilot work to develop agro advisory system using ontology. This system will be also a model for building agro advisory system for other crop which leads to increase in the production of crop.

REFERENCES

1. Agrisnet(2014).A Mission mode project to promote agricultural informatics and communications, http://agrisnetodisha.ori.nic.in/ June, 2014.

2. Agropedia(2014) A free and open source Java framework for building Semantic Web and Linked Data applications. http://agropedia.iitk.ac.in/, June,2014.

3. Apache Jena(2014). A free and open source Java framework for building Semantic Web and Linked Data applications. http://jena.apache.org, June, 2014.

4. Aree Thunkijjanukij(2009).*Ontology development for agricultural research knowledge management: a case study for thai rice*, PhD thesis, Graduate School, Kasetsart University, 2009.

5. Arun Pande, Bhushan G. Jagyasi, Sanjay Kimbahune, Pankaj Doke, Ajay Mittal, Dineshkumar Singh, and Ramesh Jain(2009).*Mobile Phone based Agro-Advisory System for Agricultural Challenges in Rural India(mKrishi)*, in IEEE Conference on Technology for Humanitarian Challenges, Aug 2009.

6. G.K. Koutu,P.P Shastry, D.K. Mishra, and K.C. Mandloi(2014).*Handbook of COTTON*, Studium Press Pvt. Ltd,India, 2014.

7. KissanKerela(2014) An Integrated multi-modal agricultural information system for kerela," IIITM-K, Thiruvanathapuram, http://www.kissankerala.net.June.2014.

8. M. H. Bohara, M. Mishr, and S. Chaudhary(2013).*RESTful Web Service Integration Using Android Platform*, in Fourth International Conference on Computing Communication and Networking Technologies , Tiruchengode, Tamilnadu,India, July 4-6, 2013.

9. M.Sabesh(2014),CICR: Approved Package of Practices for cotton," http://www.cicr.org.in/pop/gj.pdf, June, 2014. [20]

10. P. Krishna Reddy, G. V. Ramaraju, and G. S. Reddy(2007)."*eSagu™: a data warehouse enabled personalized agricultural advisory system* in SIGMOD'07 Proceedings of the 2007ACM SIGMOD international conference on Management of data, pp.910-914,June 2007.

11. OWL: Web Ontology Language(2014). http://www.w3.org/2001/sw/wiki/OWL, June, 2014

12. Ramamritham Krithi, Anil Bahuman, Ruchi Kumar, Aditya Chand, Subhasri Duttagupta, G.V. Raja Kumar, and Chaitra Rao(2004).*aAQUA Multilingual, Multimedia Forum for the community*, in IEEE International Conference on Multimedia and Expo, vol. 3,2004.

13. RanjanaAgrawal(2012).http://www.iasri.res.in/ebook/fet/Chap%205_Weather%20based%20preharvest%20crop%20forecasting.pdf

14. RDF: Resource Description Framework(2014). http://www.w3.org/RDF/, June, 2014.

15. S.Chaudhary and M. Bhise(2013).*RESTful Services for Agricultural Recommendation System* in Proceedings of NSDI-2013, IIT, Mumbai,pp. 46-52, November 29-30,2013).

16. S.Sahni(2012).*Ontology Based Agro Advisory System*, Department of Computer Science and Engineering, IIT Mumbai, M.Tech. thesis , June 2012.

17. SPARQL:SPARQL Protocol and Query Language for RDF(2014). http://www.w3 .0rgITRlrdf-sparq I-query/, June, 2014.

18. Thomas B. Freeland Jr. (lead), Gordon Andrews, Bill Pettigrew, and Peggy Thaxton.Guide to Agricultural Metrological Practices, Chapter 10.1 *AGROMETEOROLOGY AND COTTON PRODUCTION* ", World Meteorological Organization publication, Switzerland, ISBN 978-92-63-10134-1.

19. Tim Berners-Lee, James Hendler and OraLassila(2001).The Semantic Web, *Scientific American*, pp. 02, May 17, 2001.

20. V.Kumar, V. Dave, R. Nagrani, S. Chaudhary, and M. Bhise(2013).*Crop Cultivation Information System on Mobile Phones*, in IEEE Global Humanitarian Technology Conference (GHTC), pp. 196-202, ISBN: 978-1-4799-1094-6,2013.

21. V.S.Nagare(2012) CICR: Cotton plant health. http://www.cicr.org.in/pdf/cotton_plant_health.pdf

22. XML: Extensible Markup Language,(2014). http://www.w3.org/XMLI, June, 2014.