

Article

Knowledge-Based System for Crop Pests and Diseases Recognition

Miguel Ángel Rodríguez-García ¹, Francisco García-Sánchez ^{2,*}  and Rafael Valencia-García ² 

¹ Department of Computer Science, Rey Juan Carlos University, 28933 Madrid, Spain; miguel.rodriguez@urjc.es

² Department of Informatics and Systems, Faculty of Computer Science, University of Murcia, 30100 Murcia, Spain; valencia@um.es

* Correspondence: frgarcia@um.es; Tel.: +34-868-88-8107

Abstract: With the rapid increase in the world's population, there is an ever-growing need for a sustainable food supply. Agriculture is one of the pillars for worldwide food provisioning, with fruits and vegetables being essential for a healthy diet. However, in the last few years the worldwide dispersion of virulent plant pests and diseases has caused significant decreases in the yield and quality of crops, in particular fruit, cereal and vegetables. Climate change and the intensification of global trade flows further accentuate the issue. Integrated Pest Management (IPM) is an approach to pest control that aims at maintaining pest insects at tolerable levels, keeping pest populations below an economic injury level. Under these circumstances, the early identification of pests and diseases becomes crucial. In this work, we present the first step towards a fully fledged, semantically enhanced decision support system for IPM. **The ultimate goal is to build a complete agricultural knowledge base by gathering data from multiple, heterogeneous sources and to develop a system to assist farmers in decision making concerning the control of pests and diseases.** The pest classifier framework has been evaluated in a simulated environment, obtaining an aggregated accuracy of 98.8%.



Citation: Rodríguez-García, M.Á.; García-Sánchez, F.; Valencia-García, R. Knowledge-Based System for Crop Pests and Diseases Recognition. *Electronics* **2021**, *10*, 905. <https://doi.org/10.3390/electronics10080905>

Academic Editor: Rui Pedro Lopes

Received: 6 March 2021

Accepted: 7 April 2021

Published: 10 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: agrisemantics; crop pest recognition; natural language processing; ontology population; semantic data integration

1. Introduction

The World Health Organization (WHO) and the Food and Agriculture Organization (FAO) of the United Nations agreed on the following definition: *organic agriculture is a holistic production management system which promotes and enhances agroecosystem health, including biodiversity, biological cycles, and soil biological activity. It emphasizes the use of management practices in preference to the use of off-farm inputs, taking into account that regional conditions require locally adapted systems. This is accomplished by using, where possible, cultural, biological and mechanical methods, as opposed to using synthetic materials, to fulfil any specific function within the system* [1]. Thus, beyond ensuring the provision of food for the increasing world population, organic agriculture is concerned with sustainability [2]. If pests and diseases are one of the main threats to crop yields when employing conventional farming practices, in organic agriculture, in which the application of synthetic chemical fertilizers and pesticides is prohibited, the impact could be devastating [3]. For that reason, the general approach in organic agriculture is to apply management practices aiming at preventing pests and diseases from affecting a crop, rather than treating the symptoms. On the side, globalization and climate change are contributing to the emergence of new diseases and to their spread [4,5]. Under these circumstances, early detection of the outbreak of a pest or disease becomes paramount to reduce yield losses and their corresponding economic damage. Both small and large farm owners should be provided with access to relevant information about best practices in organic agriculture and the allowed methods to fight crop pests and diseases. However, in most cases such information is dispersed throughout

multiple, heterogeneous data sources. Information and Communication Technology (ICT) should play an important role in attaining such an objective.

As with many other fields, the adoption of ICT in agriculture, also known as e-agriculture, is pervasive worldwide [6]. Smart farming is the term coined to describe the application of advanced information technologies (IT) to make agriculture more efficient and effective [7]. The main applications of ICT in agriculture include the use of GPS and Geographic Information Systems (GIS) for precision farming, smartphone apps for e-learning (i.e., agroadvisory services) and crop management, RFID (Radio-Frequency Identification) for product tracking, etc. Apps in the agronomy domain can be classified in the following categories [8]: weather, soil preparation, sowing scheduling, farm management, soil fertility and crop nutrition, pest management, irrigation and drainage, precision agriculture, teaching and research apps. The number of ICT-enabled solutions for the recognition of plant pests and diseases is ever increasing [9,10]. While most of the proposed tools are based on image processing [11], other approaches such as natural language-guided, rule-based engines have also been explored [12]. The majority of the existing applications provide guidance to farmers as to how to control the outbreak of the identified pest. However, instead of building upon the vast amount of information already available about the topic, a manual process involving agronomy experts is usually required to prepare the knowledge base. This is partially due to the difficulty associated with the integration of heterogeneous data coming from disparate sources. Ontologies and other related semantic technologies have proven effective for data integration in multiple domains [13–15].

The Semantic Web adds semantics to the data published on the Web (i.e., defines the meaning of the data), so that machines are able to process these data in a similar way to what a human can do [16]. The logical formalisms behind ontological models allow autonomous agents to interpret the information that is being processed [17]. They also facilitate the execution of reasoning and inferring processes over these data. A number of different tools that make use of semantic technologies to improve classifiers and recommender systems have been developed in the last few years [18,19]. Semantically enhanced tools have also been built in the field of agronomy [20]. Agrisemantics (<http://agrisemantics.org/>, accessed on 9 April 2021) is an initiative intended to foster the use of formal semantics to enable the interoperability of agricultural data. In this work, an ontology-based expert system (we characterize the proposed framework as an “expert system” for disease diagnosis and treatment recommendations since it is a computerized system that emulates the decision-making ability of a human expert) for crop pests and diseases recognition is proposed as a major extension of our previous work [21]. The aim of the suggested approach is to enable the integration of agricultural data from heterogeneous sources. This is done by employing the CropPestO ontology [22], an ontological model in the crop pests and diseases domain that has been automatically populated from unstructured documents by leveraging natural language processing (NLP) techniques. The source documents are the official Spanish guides on crop pests and plant pathogens, and their control with Integrated Pest Management (IPM) practices, that aim to keep the use of control methods to levels that are economically justified [23]. The resulting knowledge base contains relevant information about known pests and diseases, their most common symptoms, and the suggested treatment both for conventional and organic agriculture. Once this knowledge base is reachable, our classifier (we characterize the proposed framework as a “classifier” since it can be used to associate a set of symptoms -inserted by means of sentences in natural language- with pests; the pests or diseases related to a given crop are the classes into which the set of entered symptoms can be classified) framework can readily determine the pest or disease that is most likely present in the crop given a list of symptoms expressed in natural language (i.e., the observations made by the farmer in situ). A further contribution of our work is that the focus is mainly set on organic agriculture practices and permitted pest control measures, thus disseminating organic agriculture policies and encouraging these sustainable agricultural farming practices. The proposed framework should serve as the

core for a fully fledged pest recognition system in which other evidentiary inputs, such as images of the damage and environmental parameters (e.g., weather, soil condition, etc.), are considered to boost the effectiveness of the system.

The main contributions of this work can be summarized as follows:

- A knowledge model representing the crop pest domain in the form of an ontology has been defined as a revision of our previous work [22].
- Natural language processing tools have been used to automatically populate the ontology from unstructured data sources in Spanish. During this process, data about the plant pests and diseases, including their symptoms and recommended treatments, are gathered.
- A novel approach for representing symptoms associated plant diseases is proposed based on the combination of plant parts and observed damage.
- A knowledge-based crop pest recognizer has been built which is capable of identifying multiple overlapping pests. The proposed framework is easily extensible to support new evidentiary inputs such as images.
- Sustainable agricultural practices are fostered by suggesting organic agriculture-compliant treatments. Nonetheless, the proposed framework provides support for both conventional and organic management strategies.
- A dataset has been compiled with symptoms–pests associations for three crops: almond tree, olive tree and grape vine. In total, the dataset contains 212 symptoms declared by means of sentences in Spanish, connected to 75 pests and diseases. This dataset is publicly available at <http://agrisemantics.inf.um.es/datasets/> (accessed on 9 April 2021).

The rest of this paper is organized as follows. Section 2 provides the background information on IT-enabled tools for plant pests and diseases recognition and current approaches to knowledge acquisition from natural language texts. The framework proposed in this work to automatically identify the pest or disease that has affected a crop is described in Section 3. Section 4 shows a preliminary validation analysis of the framework in a simulated environment, and finally, our conclusions and future work are put forward in Section 5.

2. Related Work

In this paper, a plant pest diagnosis system is described that leverages an automatically populated knowledge base to boost the overall accuracy results. In the last few years, researchers in the field of agronomy have proposed a significant number of ways of recognizing plant pests and diseases, while semantic technologies have simultaneously been leveraged to improve the performance of natural language processing tools in different application domains. In this section, various approaches to plant diseases' identification and management will be discussed and the most representative works in ontology-driven natural language processing will be listed.

2.1. Pests and Diseases Recognition

Among the uses of ICTs in agriculture, that of automatic pests and diseases recognition is extensive [24,25]. The most common approach is that of image processing [11,26,27] using sophisticated artificial intelligence techniques such as deep learning [28–34]. In some cases, image processing is complemented with information retrieved by sensors [35–37] or other inputs [38]. Scarcer are the approaches relying on other evidence, such as odor [39,40], weather [41,42], or rule-based systems triggered by symptoms introduced manually in natural language [43–47]. While some solutions focus on a specific crop or a single condition (throughout this manuscript the term “condition” is used as a synonym for “pest or disease”) reaching very high accuracy values [35,41,48–53], others struggle to achieve good precision results dealing with a large number of conditions in different crops [38,54–56]. A common issue hampering image-based tools for plant pest identification is that of the scarcity of images available to train the deep learning method in question [57]. To overcome

this limitation, in Reference [58] the authors present a method to generate complete plant lesion leaf images with the aim to assist in improving the recognition accuracy of the classification tool.

In previous works, we have recently explored different syntactic-based approaches to pest identification [21,59,60]. In Reference [59] we describe a tool that relies on Google Cloud's Vision API (<https://cloud.google.com/vision/>, accessed on 9 April 2021) to recognize the pest or disease affecting a plant from an image taken by the farmer. The image is sent to the Vision API, which automatically assigns labels to the image. The labels are then compared against the data available in a local database containing information about common diseases, their symptoms and suggested treatments. If one of the labels matches the name of one of the recorded pests or diseases, then all the information about that pest or disease is shown to the user. Similarly, in Reference [60] an image-based approach is proposed. Here, instead of relying on an external service such as the Google Cloud's Vision API, we built our own image processing tool by using Convolutional Neural Networks. The network was trained using a large set of images about common conditions affecting stone and citrus fruit trees along with some other crops. In such "ideal" conditions (only two diseases are considered), precision reached 90%. Again, if the diagnosis is successful, both the identified condition and the suggested treatment are shown to the user. Finally, a NLP-based approach was presented in [21], where the GATE framework (<https://gate.ac.uk/>, accessed on 9 April 2021) ("General Architecture for Text Engineering") is leveraged to process the textual description of the visible symptoms and impacts of the pest or disease. The keywords retrieved by GATE are then compared against a pest control database and the most likely causes of the problem are shown along with the recommended treatments.

Image-based pest recognition techniques benefit from the current widespread use of phones with high-resolution cameras. However, the range of pests and diseases covered by these approaches is limited to those with a visible impact on plants and their structural components (stems, leaves, flowers or fruit). Other conditions associated with injury that cannot be captured by a photographic shot will not be identified by these tools (e.g., premature fruit drop). Language-guided approaches are more flexible in terms of coverage, but their accuracy is hampered by the inherent ambiguity and imprecision of natural language. Ontologies and other related semantic technologies have proven to be useful to limit the effects of language ambiguity in different scenarios [18,61,62]. The use of knowledge technologies in agriculture is very broad [20]. Currently, there are several ontologies and structured vocabularies available in the agronomy domain [63–67]. AgroPortal [68] has become the reference repository in which most of the vocabularies and ontologies produced to represent and annotate agronomic data are hosted. More specifically, in Reference [69] the authors provide a detailed review on the use of knowledge graphs in the crop pests and diseases domain.

By building upon such formal collection of terms, several applications have been developed to assist farmers in their day-to-day practices, including pest control [12,43,70–72]. In Reference [43] the authors describe a knowledge-based system to support the diagnosis of plant diseases. The system rests on a rule-based engine built with the assistance of domain experts. If the symptoms described by the farmer trigger a rule, then a diagnosis is provided, and relevant treatments and recommendations are suggested to the farmer. The way in which symptoms are entered into the system is not clear, but it relies on the perception of the farmer. Ontologies are also leveraged in [70] to model the interrelation between crops, pests and treatments. Once the model has been automatically populated from a number of different heterogeneous sources (official guides) including 462 crops, 549 pests and 42,397 treatments, a recommendation system suggests the required treatment given the crop and the symptoms. While the approach is similar to ours (i.e., use of natural language processing to build a knowledge base with which to nourish a recommendations system), the focus is set on different stages of the process. In our work, the main goal is to assist farmers in identifying the pests and diseases in their crops; meanwhile, in

Reference [70] the authors put their attention on obtaining a fully fledged knowledge base, undermining the symptoms–pest matching. The symptoms in our work are modelled in the form “plant part-damage”, thus simplifying the matching between the symptoms in the knowledge base associated with each pest and the symptoms entered by users (which are also processed using NLP). In Reference [70] a text field is used to store the textual description of the produced symptoms; then a regular expression with the symptoms indicated by users is embedded in a SPARQL query for finding matches in the knowledge base. Additionally, one of the main drawbacks of this work is that it has not been validated using traditional information retrieval or recommender systems evaluation metrics. The authors in [71] present an ontology-based agro advisory system with the aim to bridge the gap between farmers and agriculture domain experts by integrating various data resources. A built-from-scratch cotton crop ontology constitutes the knowledge base for the proposed expert system, which provides advice to farmers given keyword-based queries. The collection of information regarding cotton farming practices to nourish the knowledge base was done manually from different sources and with the assistance of experts in the field. Another related approach is suggested in [72], where plant diseases are modelled in the form of ontology elements and the diseases likely affecting a crop are retrieved given farmers’ observations. To issue queries to the knowledge base, those observations should be transformed into Web Ontology Language (OWL) concepts. Besides, the validation of the proposed system is limited to the conditions associated with a single crop, namely, rice, for which the authors developed a rice disease ontology. Finally, in Reference [12] the authors describe AgriEnt, a knowledge-based Web platform for assisting farmers in the crop insect pest diagnosis and management. The AgriEnt-Ontology constitutes the cornerstone of the platform, an ontology representing knowledge about crops, diseases, symptoms, insects, insect pests, and treatment recommendations. To populate the knowledge base, crop insect pests’ records generated by agricultural entomology experts as well as academic publications were collected. Then, a rule-based inference engine built using the Semantic Web Rule Language (<https://www.w3.org/Submission/SWRL/>, accessed on 9 April 2021) (SWRL) is used to explore the symptoms and provide a diagnosis. Again, experts in the field were required to define the rules. Furthermore, a diagnosis is only reached when all the symptoms defined in the rule have been pointed out by the user. The average accuracy obtained by the system for the six crops considered, namely, sugar, cocoa, corn, rice, banana and soya, is above 82%.

In this work, a novel approach to the recognition of crop pests and diseases based on the combination of language technologies and semantic conceptual representations is proposed. To build this expert system, no human expert was required since all the required knowledge was gathered from available resources. Our framework makes use of a formula to calculate the likelihood that each pest connected to a given crop is the one associated with the symptoms pointed out by the farmer. The obtained scores allow the system to provide a ranked list of the possible conditions affecting a crop. As a consequence, if more than one pest or disease is actually present, the farmer can become aware of such a circumstance.

2.2. Language Technologies for Knowledge Acquisition

The manual construction of ontologies is a demanding task which needs a great deal of time and resources. To avoid it, several studies have been conducted lately on their automatic construction and update [73,74]. It is possible to distinguish three main categories: ontology learning, ontology population (a.k.a., ontology instantiation), and ontology evolution (a.k.a., ontology enrichment). Ontology learning involves the extraction of new concepts, relations, attributes, and axioms [75,76]. Due to this processing, the terminological component of ontologies (TBox) is modified. On the other hand, the automatic instantiation of ontologies [77] extracts and classifies the instances of the concepts and features which have been defined by ontologies (ABox). The starting point of the ontology’s instantiation is usually a partially instantiated ontology or a combination of

possible individuals or named entities and relations between those entities. The stages of ontology learning and instantiation from text in natural language are mainly term extraction, synonym detection, concept creation, named entity detection, the creation of concept hierarchies, the extraction of other nontaxonomic relations, and axiom acquisition [78]. Although the stages for the creation of concept hierarchies have obtained very good results for different languages, further research is currently being conducted on the automatic extraction of taxonomic relations, nontaxonomic relations and axioms [79], since current approaches have not yielded satisfying results [80]. The main problem of these automatic extraction strategies is that most of them aim to detect a predetermined combination of relations such as partonomy, time, causality, etc. [79]. As regards axiom extraction, there are few studies trying to extract simple axioms like those dealing with nontaxonomic relations [80,81]. Another drawback is that there is limited research to fulfill this task in Spanish, where the authors have made a major contribution [82]. The evolution of ontologies is based on the two technologies explained above and it not only deals with the creation of new information, but also with the updating (creation, modification and deletion) of elements as the domain changes over time. Currently, there are some satisfactory solutions [83], but they pose the same problems that those mentioned above for language technologies and automatic instantiation of ontologies.

On the other hand, text annotation can be considered as a process that enables the mapping of concepts, relations, comments or descriptions to a document or a text extract. Overall, annotations can be assimilated to metadata associated with particular text extracts from a document or any other pieces of information. Semantic annotation helps deal with natural language ambiguity and its representation through ontologies [84,85]. This process involves relating text extracts with tags representing ontological elements (concepts, relations, attributes, and instances), which enables document processing by software systems. The major limitation of these methodologies is their reliance on static knowledge; thus, the ontologies do not evolve over time. Recent studies conducted by the authors of this work provide tools for semantic annotation based on ontology evolution technology [86,87]. Finally, it is worth noting that new deep learning technologies are being applied to traditional ontology learning tasks in different languages [88].

In this work, we built upon existing natural language processing resources to develop an automatic ontology population tool which is used to gather relevant data from unstructured documents and create the corresponding instances in the ontology. For future work, we plan to exploit our previous experience in ontology evolution to apply refinement actions and enable the adaptation of the knowledge base to this changing domain.

3. Crop Pests and Diseases Identification from Natural Language Text

In this work, an expert system to classify symptoms expressed in natural language into crop pests and diseases is proposed. This section provides a detailed description of the proposed framework. Next, the functional architecture of the framework is presented, and its main components are explained.

3.1. Proposed Framework

The functional architecture of the proposed system is shown in Figure 1 and comprises three main components: (i) the pests and diseases management ontology (CropPestO); (ii) the knowledge base population tool (KB Instantiator); and (iii) the crop symptoms analyzer. The input to the system is a list of symptoms expressed in natural language that represent the harmful effects of a likely pest or disease affecting a given plant (users select the crop from a list of the crops found in the knowledge base), while the output is an ordered list of crop pests and diseases matching the provided symptoms.

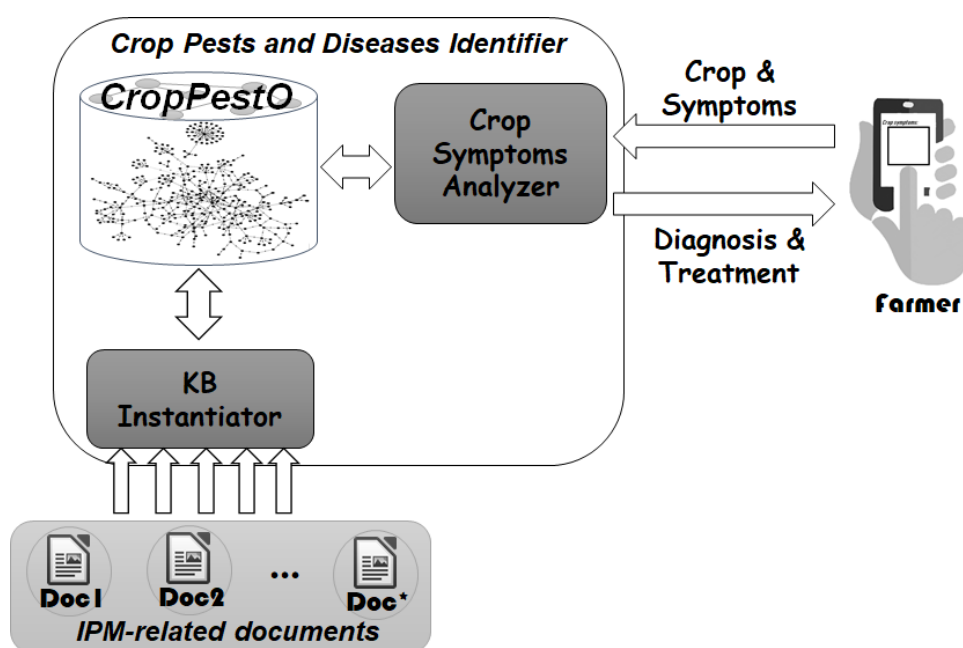


Figure 1. Proposed functional architecture framework for pests and diseases identification from natural language text.

The system works as follows using a two-step process. The first step takes place before the system is made available to users and consists of the population of the knowledge base. During this stage, a number of unstructured documents elaborated by experts in the field of IPM are processed by the KB Instantiator. This component takes into account the reference data model, a previously defined pests and diseases management ontology named CropPestO, to transform the natural language input into a number of instances to be added to the knowledge base. Once the knowledge base has been fully populated, the system becomes functional and users can interact with it. Users should point out both the crops that are being likely affected by some condition (that is, users select the crop under question from the list of crops included in the knowledge base) and the observed symptoms, which are defined in natural language. During this second stage, the Crop Symptoms Analyzer processes the entered symptoms and matches them with the ones previously introduced in the knowledge base. From the matches found, a ranked list with the most likely conditions along with the suggested control methods are shown to the users.

3.2. CropPestO: Pests and Diseases Management Ontology

The pests and diseases knowledge base, which constitutes the cornerstone of the proposed approach, is based on a domain ontological scheme that has been designed by following the steps suggested in the “Ontology Development 101” guide [89]. While there are a number of ontologies in the agronomy domain and, more specifically, in the crop pests and diseases field, none fit the requirements of an organic agriculture-based pest control recommender system. The scope of the ontology has been limited by the following competency questions (i.e., questions the ontology should help to answer): (i) Which measures should be applied to prevent the outbreak of a disease or pest? (ii) What evidence does an outbreak of a disease or pest suggest in a crop?; (iii) Which disease or pest is present in a crop?; and (iv) Which measures should be applied at any given moment to treat a disease or pest? The focus is set on organic agriculture, so organic-compliant control methods are highlighted.

In the development of the ontology, some of the terms included in the AGROVOC thesaurus [90] were reused. AGROVOC is a controlled vocabulary built by United Nations’ FAO, with more than 37,000 concepts and 750,000 terms in up to 37 languages covering elements related to food, nutrition, environment, plant cultivation techniques, etc.

AGROVOC was just recently published as a linked open data (LOD) set and is aligned with other 18 datasets related to agriculture. Besides, AGROVOC satisfies our requirements in terms of both completeness (including a large number of domain relevant concepts and being actively maintained and updated (<http://aims.fao.org/agrovoc/releases>, accessed on 9 April 2021); in particular, AGROVOC includes concepts tagged in both English and Spanish for all the pathogens, plants and plant products covered in the processed IPM-related documents) and formality (enough semantic expressivity for our purposes). The upper-level concepts that form the backbone of the ontology are as follows: “Plant Product” (i.e., a product produced by a plant including cereals, fruits, legumes, among others), “Pest” (i.e., this concept encompasses both diseases and pests that can inflict damages on plants or plant products, such as fruit flies or tuta absoluta), “Control Method” (i.e., technique that can be applied to avoid or reduce the harmful effects of pests and diseases, such as trap cropping or sexual confusion), “Plantae” (i.e., plant; the focus is set on those that produce basic human foods such as grain, fruit and vegetables, including *Solanum lycopersicum* or *Prunus armeniaca*, among others), “Symptom” (i.e., a physical feature which is regarded as indicating a condition of disease, such as fruit rot or leaf spot). Additionally, to assist in the resolution of the abovementioned competency questions, the following relationships between the upper-level elements were considered (along with their inverse relationships): “Plantae produces Plant Product”, “Plant Product hasPest Pest”, “Symptom isInfluencedBy Pest”, and “Control Method controls Pest”.

The ontology has been developed in OWL 2 [91] and is available at <http://agrisemantics.inf.um.es/ontologies/CropPestOv2.owl> (accessed on 9 April 2021). More details on the ontology construction process can be found in [22]. An excerpt of the ontology, including the high-level classes, is depicted in Figure 2. In the figure, the classes and relationships directly extracted from AGROVOC are represented in green. Since the framework has been originally conceived to be used by Spanish-speaking farmers (e.g., the documents used for populating the ontology are Spanish reference guides for IPM in different crops, see Section 3.3), the ontology has been labelled in Spanish (besides English). In total, the populated ontology contains 286 classes, 8 object properties, 11,754 individuals, and 96,550 axioms. The correctness of the resulting ontology has been checked using the following tools: (i) the RDFS Validator (<https://www.w3.org/RDF/Validator/>, accessed on 9 April 2021) to repair the definitions of concepts, relations, and instances; (ii) OOPS! (<http://oops.linkeddata.es/catalogue.jsp>, accessed on 9 April 2021) (the Ontology Pitfall Scanner!) to identify deficiencies in metadata information such as license and version information, among others; and (iii) OQuare (<https://semantics.inf.um.es/ontology-metrics/>, accessed on 9 April 2021) to test the model’s features.

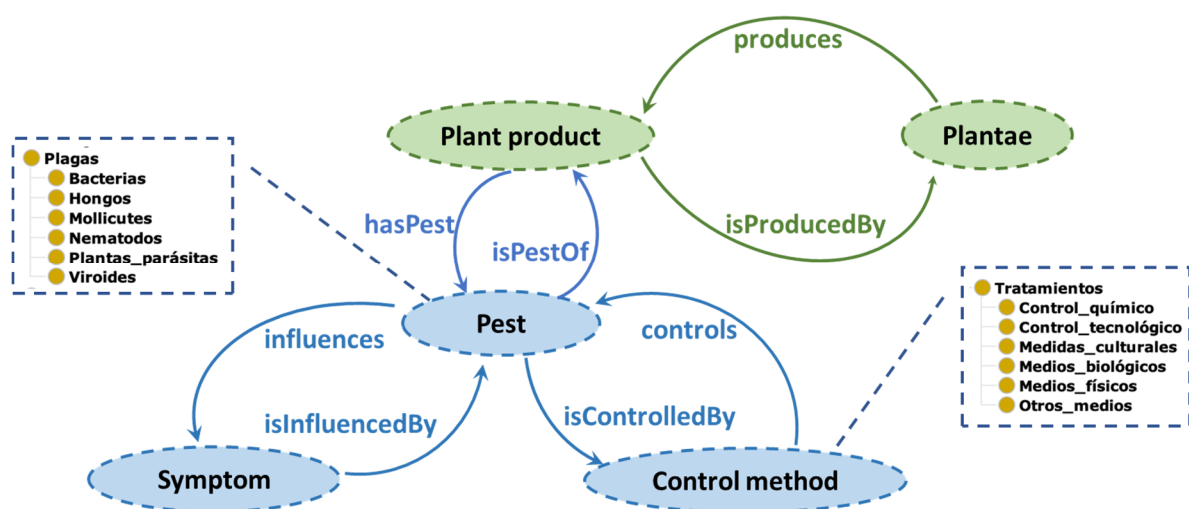


Figure 2. Ontology’s high-level classes and their relationships (partially in Spanish).

3.3. KB Instantiator: Knowledge Base Population

In agriculture, pest control is a vast field where each crop can be infected with different types of infectious agents. In this work, the focus is set on the crops grown in Spain, but the proposed framework can be easily adapted to other environments. To prepare the list of supported crops, the Web portal of the Spanish Ministry for Agriculture, Fisheries and Food (<https://www.mapa.gob.es/es/agricultura/temas/default.aspx>, accessed on 9 April 2021) has been queried. It contains detailed information and statistics about the production and exploitation of the crops in this country. Besides, it offers a set of official documents written by agronomy experts to guide farmers in applying the most convenient control methods according to an IPM strategy to combat the pests and diseases that affect a number of different crops (<https://www.mapa.gob.es/es/agricultura/temas/sanidad-vegetal/productos-fitosanitarios/guias-gestion-plagas/>, accessed on 9 April 2021). Each document provides information about the pathogenic agents known to attack a given crop. It also describes their associated symptomatology and provides details about the most appropriate strategies to monitor, prevent and combat these infectious agents. These IPM-related documents have been used to create the relationships between crops (i.e., “Plant Product”) and their associated pests and diseases (i.e., “Pest”), and between each pest/disease (i.e., “Pest”) and the associated symptoms (i.e., “Symptom”) and the suggested treatment (i.e., “Control Method”), as described below. In this section, the process carried out to populate the ontology is described in detail. This process encompasses the steps enumerated next.

A variety of dictionaries have been created to assist in both the definition of the ontology scheme and in its instantiation. The process starts with the definition of the list of supported crops. To create this list, the report available in [92], which studies the performance of crops and crop groups with great relevancy in the Spanish economy, has been analyzed. The document includes an annex with a list of the crops produced in towns and provinces. The Snowtide library (<https://www.snowtide.com>, accessed on 9 April 2021) has been used to process the PDF file, extract the list of crops and create the glossary of crops. Besides, the resulting list has been manually enriched with the plants that produce the crop. The resulting resource is a file with a list of pairs relating each crop (i.e., “Plant Product”) with the plant producing it (i.e., “Plantae”).

Similarly, the IPM-related documents mentioned above have been processed to extract: (i) pest names (i.e., “Pest”); (ii) the damages produced by such pests (i.e., “Symptoms”); and (iii) their recommended treatments (i.e., “Control Method”). As a result, for each document two new files are defined, one including pest names, the crops known to be attacked by such pests and diseases, and the associated symptoms and damages produced, and the other associating each pest and the recommended control methods, which are described in tabular format. In addition to this, additional information was gathered from the resource at [93]. This official document provides a detailed description of a wide variety of pathogen agents of plants. It offers a complete classification of the plant pathogens observed in Spain including virus, viroid, bacteria and fungus, among others. It also provides specific sections with further details about synonyms, the taxonomy they belong to, associated symptoms, and hosts affected. A script was implemented to process each pathogen from the document and extract the associated details. In all the processed documents, the connection between pests, the plant products that they harm, their symptomatology, and the known control methods to limit their impact and spread are made explicit and can be easily reproduced in the knowledge base.

However, to facilitate the search for symptoms in the knowledge base matching those expressed by the users of the system (a key step in the pest recognition process) we conceived a novel approach to represent the symptoms. In particular, a two-step method has been defined to process the natural language sentences describing the pests associated symptomatology in the documents. First, all relevant information is gathered, and the text is tokenized into sentences—that is, the text is divided in sentences. Then, these sentences are analyzed and only those providing specific details about the effects

of the pest or disease in the plant and in the plant product are kept. To do so, a recursive method was designed which analyzes the syntactic dependency graph of each sentence. During this search, it utilizes a morphology glossary and a phytopathology dictionary (<http://www.ub.edu/vocabularia/archives/4518>, accessed on 9 April 2021) to identify terms related to part of plants and damages, respectively. The morphology glossary was created in a semisupervised manner. First, a frequency analysis tool was used to identify the most used words in the sections dealing with pests' symptomatology. A stopwords list was employed next to remove nondomain specific terms. Then, a botanical dictionary (<https://www.arbolesornamentales.es/glosario.htm>, accessed on 9 April 2021) was exploited to identify those words specifically related to the plants' domain. As a result, a list of terms sorted by the number of apparitions is obtained. Finally, the list was analyzed to remove those identified words which are related to the plant domain but are not explicitly related to the plants' morphology. Once the relevant sentences have been identified, they are processed so that "plant part–damage" pairs are obtained, generating a new resource in which each pest is associated with the gathered pairs. Algorithm 1 describes the pseudocode of the two-step method.

Algorithm 1: SymptomsAnalyser

```

Function analyseGraph(root, stack):
    if token in sw_dict then
        | return;
    if token in fito_dict then
        | stack ← token;
    for children in token.children do
        | analyseGraph(child,stack);

morfo_dict ← load(mg);
fito_dict ← load(dp);
sw_dict ← load(sw);
tokens_dict ← load(s);
while tokens ≠ ∅ do
    | token ← tokens if token in sw_dict then
    | | continue;
    if token in morfo_dict then
    | | root ← getRootSentence();
    | | analyseGraph(root, stack);
    | | write(stack);

```

As inputs, the method receives four parameters: *s*, *mg*, *dp* and *sw*. The *s* parameter stands for the sentences to be analyzed; the *mg* represents the morphology dictionary utilized to identify parts of the plant; the *dp* parameter denotes the phytopathology dictionary to identify the plants' injuries; and finally, the *sw* indicates the stopwords list used to filter terms without useful information such as prepositions and conjunctions, among others. The method utilizes the library spaCy (<https://spacy.io>, accessed on 9 April 2021) to extract the syntactic dependency graph of a sentence. SpaCy is an open-source library for Natural Language Processing capable of analyzing a vast of text volume. Among the diverse functions that it provides, it is possible to highlight Name Entity Recognition, Part-of-speech tagging, Syntax-driven sentence segmentation, integrated viewers for syntax, etc. When a sentence is given, the method splits the sentences into tokens. First, it checks if the token is a term related to a plant morphology by using the dictionary. If so, the next

step will be to traverse into the dependency graph to obtain the verb root of the sentence. Then, it traverses the graph recursively to analyze the terms related to this token. For each extracted term, the method checks if it is related to phytopathology domain by using a dictionary. If so, the term is stored in a stack. The procedure finishes when all token's dependencies are analyzed.

Once all the resources described above are available, the ontology scheme can be enriched and the knowledge base populated. Initially, the algorithm loads the aforementioned dictionaries. Then, it first creates the base taxonomy and defines the object properties. The initial structure is composed of the high-level classes as depicted in Figure 2, namely, "Plantae", "Plant Product", "Pest" (and its hierarchy), "Symptom", and "Control Method" (and its hierarchy). Then, the method starts evolving the ontology's hierarchy by inserting crops. To integrate the crops in the CropPestO ontology, an algorithm has been implemented that recursively traverses through the AGROVOC hierarchy, collecting the upper categories of a given concept. When a term referring to either a plant product or a plant is retrieved from one of the entered dictionaries, the method localizes the concept in AGROVOC and recursively iterates through upper categories, systematically adding each concept found as subclasses in the hierarchy. The process finishes when the top concept "Plantae" or "Plant Product" is found. As a result, the whole path from the given concept to those high-level classes is inserted in the CropPestO ontology. Next, it inserts the symptoms relating them with the crops and the pests. As described above, the symptoms dictionary not only contains lists of "plant part–damage" pairs; it also indicates the pest producing those symptoms and the affected. When a pair is inserted in the ontology as an instance of "Symptom", its relationship with the condition causing such effects (i.e., "Pest") is defined employing the "influences" object property, and also is the relationship between the later and the plant product (i.e., "Plant Product") afflicted by such disease through the "hasPest" object property. Finally, the method integrates the treatments (i.e., "Control Method"). As pointed out above, in the guides, each treatment is related to a particular pest. Thus, to relate this information, it is enough to look for each pest in the model and associate its respective treatment.

As a general overview of the content of the populated ontology, in Table 1. the plant products connected to the highest number of pests are enumerated. Then, in Table 2. the pests linked to the highest number of symptoms are put forward. Finally, the symptoms associated with the highest number of pests are listed in Table 3.

Table 1. Plant products and number of associated pests.

	Plant Product	#Pests
1	"Grapes"	46
2	"Peaches"	38
3	"Apples"	36
4	"Almonds"	31
5	"Cherries"	23

Table 2. Pests and number of associated symptoms.

	Pest	#Symptoms
1	"Cryphonectria parasitica"	65
2	"Armillaria mellea"	49
3	"Phomopsis actinidiae"	49
4	"Taphrina spp"	39
5	"Botryosphaeria dothidea"	37

Table 3. Symptoms and number of associated pests (partially in Spanish).

	Symptom	#Pests
1	“Hojas podredumbre”	11
2	“Flores caída”	7
3	“Ramas manchas”	5
4	“Tronco anillos”	3
5	“Troncos chancros”	2

3.4. Crop Symptoms Analyzer

The analyzer represents the input of the system, and it provides the farmer with a natural language interface to interact with it. Through this interface, a farmer can describe the symptoms observed in a determined plant (the input is composed of (i) the plant, from a list of plants supported by the system, that is, plants available in the knowledge base, and (ii) a list of observed symptoms), and the module will answer with a list of pests and diseases ranked based on matches found with the symptoms of the diseases described in the knowledge base. Thus, when the module receives a symptom description, it employs the algorithm described above (see Algorithm 1) to decompose the sentence and keep only those tokens related to the plant domain. As a result, a list of pairs will be retrieved where each pair expresses a plant part and its damage. In the next stage, the populated CropPestO ontology is utilized to find the symptoms linked to each pair. Certainly, for each pair elaborated from the user input, the analyzer tries to find instances of the class “Symptom” matching with such a pair. If a match is found, then the instances of “Pest” related to such an instance of “Symptom” are automatically selected as a candidate to be put in the pests’ recommendation list. For example, let us suppose that the farmer writes the following sentence: “The almond tree produces black fruit”. First, the module would employ the algorithm to keep those terms related to the plants’ domain. The algorithm would analyze the sentence, and it would obtain “black fruit”, “black almond” as a list of symptoms. Next, for each pair, the populated ontology would be queried to find an exact match. If such a coincidence is found, then the method selects the associated pests, and it adds them to the recommendation list of pests to be sent back to the farmer.

The formula used to rank all candidate pests given the symptoms entered by the farmer takes into account a measure of sensitivity (importance of the symptom in the total pool of symptoms associated with a given pest) and specificity (number of pests to which a given symptom is associated). The formula is as follows:

$$score(p_j, s) = \frac{\sum_{i=1}^n (sensitivity(s_i, p_j) \times specificity(s_i))}{n}, \quad (1)$$

where p_j is one of the candidate pests considered, s is the list of all the symptoms entered by the farmer, n is the total number of symptoms provided by the farmer, $sensitivity$ is calculated as follows:

$$sensitivity(s_i, p_j) = \begin{cases} \frac{1}{|symptoms(p_j)|}, & \text{if } symptom_i \in symptoms(p_j) \\ 0, & \text{if } symptom_i \notin symptoms(p_j) \end{cases}, \quad (2)$$

where $symptoms(p_j)$ returns the set of all symptoms associated with pest p_j ; and $specificity$ is calculated as follows:

$$specificity(s_i) = \frac{1}{|pests(s_i)|}, \quad (3)$$

where $pests(s_i)$ returns the set of all pests associated with symptom s_i . The score in Equation (1) is calculated for all the pests associated with the crop at hand when at least one of the entered symptoms matches one of the symptoms associated with such a pest in the knowledge base—that is, all candidate pests. The rationale behind that formula is that (i) if a few symptoms are associated with a given pest and one of these symptoms has

been entered by the user, then that pest is a very likely candidate, and (ii) if a symptom is associated with very few pests and this symptom is entered by the user, then those pests are also very likely candidates. A candidate pest is included in the recommendation list to be shown to the user if its score is above a given threshold (to be defined by the administrator). Therefore, if the plantation is afflicted by more than one pest or disease, the farmer can become aware of such a circumstance.

4. Evaluation

This section focuses on the evaluation of the pests and diseases recognition method proposed in this work. First, an exemplary scenario is described representing how the proposed tool can be accessed by its intended users. Then, some details about the dataset used for this validation experiment are put forward and the evaluation metrics are enunciated. Finally, the main results of the experiment are shown and discussed.

4.1. Exemplary Usage Scenario

In a typical usage scenario, farmers in the field would observe some worrying signs in their plantation, the likely effects of an unknown pathogen. Under these circumstances, farmers would open the “CropPestIdentifier” app and describe the observed symptoms by means of statements in natural language. Then, the system would process the data and return a list of the pests or diseases that are most probably causing such harm. The flowchart of the app is depicted in Figure 3. The following three steps are required: (i) farmers select the crop and input the observed damage in natural language sentences; (ii) the system analyzes these inputs and leverages the knowledge base to obtain a set of pests that might be producing those damages; and (iii) farmers can visualize detailed information about each retrieved pest, including the recommended treatment.

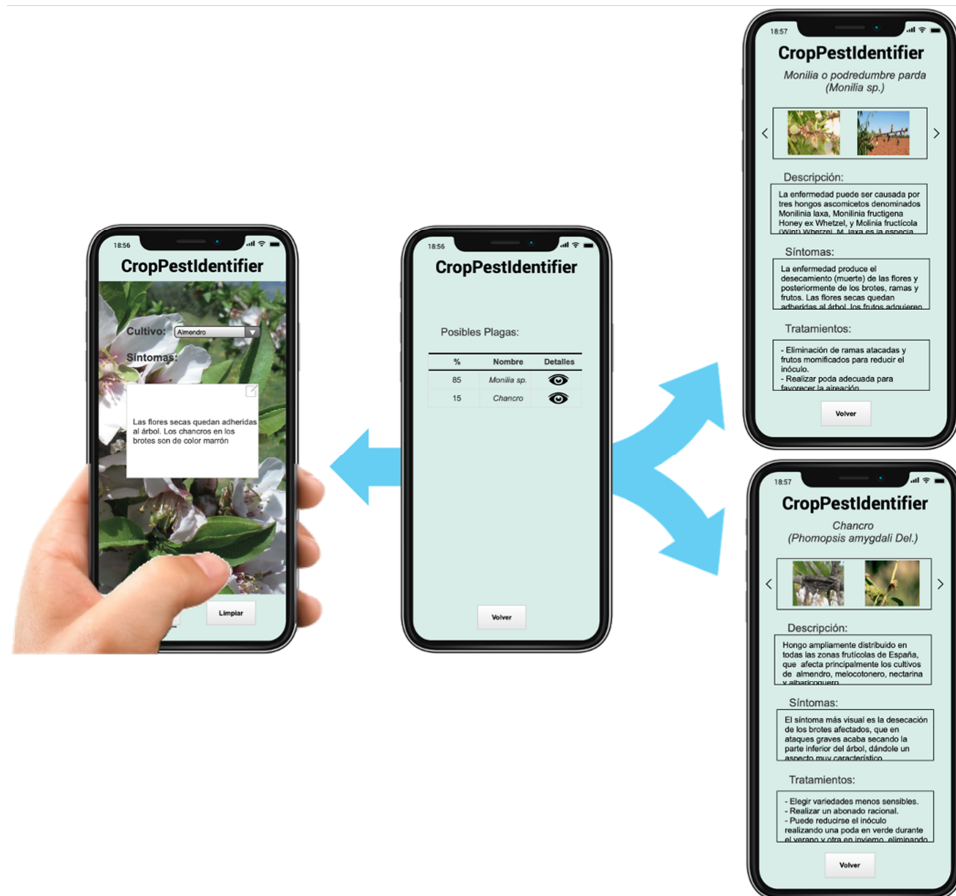


Figure 3. Exemplary use case flowchart (in Spanish).

While our “Crop Symptoms Analyzer” returns a list of the most likely pests affecting the farmer’s crops, for evaluation purposes, we only consider the pest that obtains the highest score for each test case. Consequently, it can be treated as a classification problem in which, given the crop under question and all the observed symptoms, the system has to determine the pest associated with such a crop which is more likely to be causing such symptoms. In sum, the classes into which one item of the dataset (i.e., set of symptoms) can be classified are any of the conditions (i.e., pests and diseases) associated with the crop at hand.

4.2. Dataset

For the purposes of this study, an evaluation dataset has been defined in which a number of symptoms are associated with the corresponding condition affecting a given crop. Therefore, for each pool of symptoms, the pest under question is known. The symptoms are declared by means of sentences in natural language. This dataset has been collected from a number of different webpages containing information about the pests and diseases associated with the selected crops (e.g., <https://agroes.es>, accessed on 9 April 2021, <https://www.fertibox.net>, accessed on 9 April 2021, among others) by using a web scrapping tool. An exemplary test case is shown in Table 4.

Table 4. Excerpt of the dataset (partially in Spanish).

Almond Tree Dataset
...
PEST NAME: Monilia o podredumbre parda (Monilinia sp.)
PEST SYMPTOMS AND DAMAGES:
las flores secas quedan adheridas al árbol, los frutos adquieren color negro y quedan momificados en las ramas.
los chancros en los brotes son de color marrón claro con emisiones de goma que en madera de más edad se abren.
...

In particular, the dataset built for this preliminary validation experiment contains a total of 212 symptoms, connected to 75 pests and diseases in three different crops, namely, almond tree (*Prunus dulcis*), olive tree (*Olea europaea*), and grape vine (*Vitis vinifera*). These are some of the main crops that are cultivated in Mediterranean regions. In Table 5, some additional details about this dataset are put forward. The whole test dataset is available at <http://agrisemantics.inf.um.es/datasets/> (accessed on 9 April 2021).

Table 5. Summary of the dataset content.

	Almond Tree	Olive Tree	GRAPE VINE
# of symptoms	86	42	102
# of pests	26	25	24

4.3. Evaluation Metrics

The metrics typically used to assess the performance of classification models such as the one described here are accuracy, precision, recall and f-measure. These metrics have traditionally been employed in the evaluation of information retrieval systems [94], but are well suited to the quality assessment of classifiers: we wish to verify whether the system properly identifies the pest or disease affecting the crops given some observable sign and symptoms. Four outcomes for a predicted value are consequently possible. These values are calculated for each pest in each dataset, and the results are aggregated by dataset (i.e., crop). For a given pest, (i) a True Positive (*tp*) occurs when the entered symptoms, which are associated with the pest under question, are correctly classified as being caused by this pest; (ii) a False Negative (*fn*) occurs when the pool of symptoms associated with

the pest are wrongly classified as being caused by another pest; (iii) a False Positive (fp) occurs when a pool of symptoms associated with another pest is wrongly classified as being caused by the pest under question; and (iv) a True Negative (tn) occurs when a pool of symptoms associated with another pest are not wrongly classified as being caused by the pest under question.

In this context, *Accuracy* can be interpreted as the probability of being correct and is calculated as follows:

$$Accuracy = \frac{\text{Correctly predicted}}{\text{Total number of predictions}} = \frac{tp + tn}{tp + tn + fp + fn'} \quad (4)$$

Precision, also known as positive predictive value, represents the proportion of diagnosed diseases that have been correctly classified and is obtained as follows:

$$Precision = \frac{\text{Correctly predicted as positive}}{\text{Total number of predicted as positive}} = \frac{tp}{tp + fp'} \quad (5)$$

Recall, also known as sensitivity or true positive rate, measures the system's ability to correctly classify diseases and is calculated as the proportion of actual diseases that have been correctly classified by the system:

$$Recall = \frac{\text{Correctly predicted as positive}}{\text{Total number of positives}} = \frac{tp}{tp + fn'} \quad (6)$$

Finally, the *F – measure*, also known as F1 score, is the harmonic mean of *precision* and *recall*, computed as follows:

$$F - \text{measure} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

4.4. Results

In Table 6 the results of the experiments for the four metrics considered are illustrated (more details about the results of the experiment are available at: http://agrisemantics.inf.um.es/datasets/Evaluation_results.xlsx (accessed on 9 April 2021)). The overall accuracy of the proposed approach is 98.8%, achieving a 99% accuracy for both almond tree and grape vine, and a 97% accuracy for olive tree. In a multiclass classification problem such as the one faced in this work, the precision, recall and f-measure metrics provide the evaluation on a per class basis. Given the characteristics of our dataset in which for each class (i.e., disease) only one pool of symptoms has been considered (i.e., one test for each disease), and the results are shown aggregated by crop.

Table 6. Accuracy, precision, recall and F-measure.

Experiment	Accuracy	Precision	Recall	F-Measure
Almond tree	0.993	0.846	0.885	0.859
Olive tree	0.974	0.460	0.480	0.467
Grape vine	0.997	0.813	0.958	0.819
Overall	0.988	0.707	0.773	0.716

4.5. Discussion

Generally, the classifier has achieved promising results. In the experiments carried out for both almond trees and grapes vines, only a few diseases have not been correctly classified during the experiment. Conversely, in the olive tree experiments, only half of the diseases were classified correctly, with no results for 7 of the 25 test cases. This explains the worse precision and recall values obtained, i.e., 0.460 and 0.480, with respect to 0.846 and 0.885 in the almond tree experiments and 0.813 and 0.958 in the grape vine experiments. The test cases in which no disease has been identified are those for which

our symptoms decomposition procedure could not retrieve any “plant part–damage” pair from the entered text. A more versatile approach might be required for those situations in which symptoms are not expressed as expected. On the other hand, “false positives” are usually associated with cases in which the same “plant part–damage” pair representing a symptom is linked to more than one disease. In consequence, while using common, more generally used words might result in a more human-understandable knowledge base, the overall performance of the system can be significantly degraded.

By examining the results of the evaluation process, we observed some other issues to consider. First, some of the “false negatives” (i.e., pool of symptoms not correctly associated with their corresponding disease) are due to deficiencies in the automatically generated knowledge base. Not all the symptoms pointed out in the official guides processed to populate the knowledge base have been adequately represented in the instantiated ontology. Consequently, the NLP method conceived to automatically instantiate the knowledge base should still be fine-tuned to fully tease out all relevant information. In line with this, an exhaustive analysis of the ontology model and the automatically generated instances is required to ensure the adequate representation of the original data. The method for evaluating agricultural ontologies proposed in [95] could constitute a first step towards this end.

Second, the results for diseases with few evidentiary facts are unstable. Such is the case, for example, of the *Eurytoma amygdali* Enderlein wasp in almond trees. Only one observable symptom has been identified (“black fruit”) and no matching has been found with the test input. It would be desirable to extend the pool of symptoms associated with such diseases so as to avoid this instability. Additionally, the presence of synonyms among the symptoms stored in the knowledge base and the existence of symptoms associated with more than one condition (pest or disease) can give rise to false positive results. A manual revision of the contents of this part of the ontology might become necessary. Alternatively, it is possible to simplify the way symptoms are entered in the system (and likewise represented in the ontology) by following the example of AgriEnt [12]: to present users with a list of symptoms (accompanied by representative images) from which to choose.

In the literature, the approaches closest to that presented here are AgriEnt and the information retrieval system built upon the PCT-O ontology [70]. As mentioned above, AgriEnt has been evaluated in terms of accuracy (i.e., correct diagnosis of all test cases) in six different crops, reaching an average accuracy of 0.8221. It is not appropriate to compare our respective accuracy values since they have been obtained from different experimental settings. The dataset used in the evaluation of AgriEnt has not been made public and the input is slightly different, since in AgriEnt farmers select the symptoms from a list of available symptoms associated with a given crop. On the other hand, while the population process of the PCT-O ontology and the actual ontology model are thoroughly revised in their work, the authors of this approach do not provide any performance data concerning the information retrieval or recommender system.

5. Conclusions and Future Work

Agriculture is one of the pillars for worldwide food provisioning, with fruits and vegetables being essential for a healthy diet. A large proportion of the world’s population live in countries where agriculture is the main source of livelihood [96]. Organic agriculture presents several benefits over conventional agriculture, including improved environmental health and reduction of costly external inputs [97]. However, its feasibility is often questioned due to the constraints on the use of synthetic products such as chemical fertilizers and pesticides. For that reason, the general approach in organic agriculture is to deal with the causes of a problem rather than treating the symptoms. Therefore, the early detection of a pest or disease outbreak becomes crucial so as to allow the adoption of preventive measures. Yet, in most cases farmers do not have the knowledge and resources necessary to detect the trigger factors and act accordingly. Moreover, organic agriculture-compliant treatments are still unknown to most people. It is thus necessary to provide farmers with

the means to, first, recognize the presence of pests and diseases in their crops and, second, develop preventive actions and use IPM practices allowed for organic production, to limit their harmful effects.

Many ICT-enabled tools have been developed to facilitate the detection of pests and diseases in crops. Most solutions rely on image processing and often require the use of sophisticated high-resolution image capture devices or other types of sensors that are not usually available to individuals responsible for agricultural holdings. Besides, the syntactic-based core of existing approaches limits their ability to leverage the already vast amount of information about plant pests, diseases, their causes, and their control measures. In this work, we describe a semantic approach for the identification of crop pests and diseases. The framework proposed in this paper makes use of ontologies to semantically model the domain of interest. The final knowledge base contains a total of 338 plants (i.e., individuals under the top-level “Plantae” concept) and 513 crops (i.e., individuals under the top-level “Plant Product” concept). The use of this formal model greatly facilitates the automatic integration of data from multiple, heterogeneous sources, resulting in a complete knowledge base. Reasoning and inferencing mechanisms are then put in place to determine the condition producing damages to crops and the required control measures complying with organic agriculture regulations. Actually, since IPM guides have been used to populate the knowledge base, the application can be easily extended to support conventional growers.

For future work, the CropPestO ontology will be improved to make it more human readable and incorporate more axioms to boost the inferencing capabilities. Those formal underpinnings of ontologies can then be leveraged to carry out reasoning processes that enhance pest recognition and related tasks. Besides, we plan to extend the framework to support other evidentiary items as input, including images and environmental parameters. Certainly, weather, soil conditions, affected area, affected crops, yield losses, etc., are some of the factors that can help characterize the problem’s source, and the exhaustive analysis of historic data can lead to insights into how and why certain crop pests and diseases break out. The integration of different identification systems can result in efficiency and effectiveness gains. On the other hand, currently the knowledge base has been solely populated with data from official Spanish guides, and thus is only useful for Spanish-speaking users. While the underlying ontology model has been labelled in both English and Spanish, the NLP method used for ontology population should be adapted to support other languages. Moreover, the pest control domain is an evolving, ever changing field, and so we aim to develop a semisupervised ontology evolution tool. The ontology could then be continuously enriched and updated by considering the state-of-the-art knowledge. This tool would also assist in maintaining the ontology and keeping it up to date with the changes in the reference vocabularies used. Finally, a more robust validation, in a real environment (i.e., with tests provided by real users) and with large volumes of data, is required to verify the scalability of the proposed approach. Under these circumstances the use of spell-checker tools will be essential to deal with the foreseeable typos. Synonyms should also be considered along with other matching measures such as the Levenshtein distance. As part of this envisioned validation scenario, the use of other metrics such as Mean Average Precision at k (MAP@k) and AUC-ROC (area under the ROC curve) will be considered.

Author Contributions: Conceptualization, M.Á.R.-G. and F.G.-S.; methodology, M.Á.R.-G. and R.V.-G.; software, M.Á.R.-G.; validation, M.Á.R.-G., F.G.-S. and R.V.-G.; formal analysis, R.V.-G.; investigation, M.Á.R.-G. and F.G.-S.; resources, R.V.-G.; data curation, M.Á.R.-G.; writing—original draft preparation, F.G.-S.; writing—review and editing, M.Á.R.-G., F.G.-S. and R.V.-G.; visualization, F.G.-S.; supervision, R.V.-G.; project administration, R.V.-G.; funding acquisition, F.G.-S. and R.V.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially funded by the Research Talent Attraction Program by the Comunidad de Madrid with grants references 2017-T2/TIC-5664 and Young Researchers R+D Project. Ref. M2173-SGTRS (cofunded by Rey Juan Carlos University), respectively, the Seneca Foundation—

the Regional Agency for Science and Technology of Murcia (Spain)—through project 20963/PI/18, and the Spanish National Research Agency (AEI) through project LaTe4PSP (PID2019-107652RB-I00/AEI/10.13039/501100011033).

Data Availability Statement: The populated ontology is publicly available at: <http://agrisemantics.inf.um.es/ontologies/CropPestOv2.owl> (accessed on 9 April 2021). The datasets used for evaluation are publicly available at: <http://agrisemantics.inf.um.es/datasets/> (accessed on 9 April 2021). The main results of the experiment are publicly available at: http://agrisemantics.inf.um.es/datasets/Evaluation_results.xlsx (accessed on 9 April 2021).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- World Health Organization. Food and Agriculture Organization of the United Nations. In *Organically Produced Foods*, 3rd ed.; World Health Organization: Rome, Italy, 2007; ISBN 978-92-5-105835-0.
- Reganold, J.P.; Wachter, J.M. Organic agriculture in the twenty-first century. *Nat. Plants* **2016**, *2*, 15221. [CrossRef]
- Tamm, L. The impact of pests and diseases in organic agriculture. In Proceedings of the The BCPC Conference: Pests and Diseases, Brighton, UK, 13–16 November 2000; Volume 1, pp. 159–166.
- Deutsch, C.A.; Tewksbury, J.J.; Tigchelaar, M.; Battisti, D.S.; Merrill, S.C.; Huey, R.B.; Naylor, R.L. Increase in crop losses to insect pests in a warming climate. *Science* **2018**, *361*, 916–919. [CrossRef] [PubMed]
- Velásquez, A.C.; Castroverde, C.D.M.; Yang He, S. Plant-pathogen warfare under changing climate conditions. *Curr. Biol.* **2018**, *28*, R619–R634. [CrossRef]
- Woodard, J.; Andriessen, M.; Cohen, C.; Cox, C.; Fritz, S.; Johnson, D.; Koo, J.; McLean, M.; See, L.; Speck, T.; et al. *ICT in Agriculture (Updated Edition): Connecting Smallholders to Knowledge, Networks, and Institutions*; World Bank: Washington, DC, USA, 2017; ISBN 978-1-4648-1002-2. Available online: <https://openknowledge.worldbank.org/handle/10986/27526> (accessed on 9 April 2021).
- Unal, Z. Smart farming becomes even smarter with deep learning—A bibliographical analysis. *IEEE Access* **2020**, *8*, 105587–105609. [CrossRef]
- Lagos-Ortiz, K.; Medina-Moreira, J.; Sinche-Guzmán, A.; Garzón-Goya, M.; Vergara-Lozano, V.; Valencia-García, R. Mobile applications for crops management. In Proceedings of the Technologies and Innovation—4th International Conference, CITI 2018, Guayaquil, Ecuador, 6–9 November 2018; Communications in Computer and Information Science 883. Valencia-García, R., Alcaraz-Mármol, G., del Cioppo-Morstadt, J., Vera-Lucio, N., Bucaram-Leverone, M., Eds.; Springer: Guayaquil, Ecuador, 2018; pp. 57–69. [CrossRef]
- Lee, S.H.; Goëau, H.; Bonnet, P.; Joly, A. New perspectives on plant disease characterization based on deep learning. *Comput. Electron. Agric.* **2020**, *170*, 105220. [CrossRef]
- Loey, M.; ElSawy, A.; Afify, M. Deep learning in plant diseases detection for agricultural crops: A survey. *Int. J. Serv. Sci. Manag. Eng. Technol.* **2020**, *11*, 41–58. [CrossRef]
- Sinha, A.; Shekhawat, R.S. Review of image processing approaches for detecting plant diseases. *IET Image Process.* **2020**, *14*, 1427–1439. [CrossRef]
- Lagos-Ortiz, K.; Salas-Zárate, M.D.P.; Paredes-Valverde, M.A.; García-Díaz, J.A.; Valencia-García, R. AgriEnt: A knowledge-based Web platform for managing insect pests of field crops. *Appl. Sci.* **2020**, *10*, 1040. [CrossRef]
- Bernabé-Díaz, J.A.; del Carmen Legaz-García, M.; García, J.M.; Fernández-Breis, J.T. Efficient, semantics-rich transformation and integration of large datasets. *Expert Syst. Appl.* **2019**, *133*, 198–214. [CrossRef]
- Asfand-E-Yar, M.; Ali, R. Semantic integration of heterogeneous databases of same domain using ontology. *IEEE Access* **2020**, *8*, 77903–77919. [CrossRef]
- Jiang, S.; Angarita, R.; Chiky, R.; Cormier, S.; Rousseaux, F. Towards the integration of agricultural data from heterogeneous sources: Perspectives for the French agricultural context using semantic technologies. In *Advanced Information Systems Engineering Workshops. CAiSE 2020. Lecture Notes in Business Information Processing*; Dupuy-Chessa, S., Proper, H., Eds.; Springer: Cham, Switzerland, 2020; Volume 382, pp. 89–94. [CrossRef]
- Shadbolt, N.; Berners-Lee, T.; Hall, W. The Semantic Web revisited. *IEEE Intell. Syst.* **2006**, *21*, 96–101. [CrossRef]
- Studer, R.; Benjamins, R.; Fensel, D. Knowledge engineering: Principles and methods. *Data Knowl. Eng.* **1998**, *25*, 161–197. [CrossRef]
- García-Sánchez, F.; Colomo-Palacios, R.; Valencia-García, R. A social-semantic recommender system for advertisements. *Inf. Process. Manag.* **2020**, *57*, 102153. [CrossRef]
- Shanavas, N.; Wang, H.; Lin, Z.; Hawe, G. Ontology-based enriched concept graphs for medical document classification. *Inf. Sci.* **2020**, *525*, 172–181. [CrossRef]
- Drury, B.; Fernandes, R.; Moura, M.-F.; de Andrade Lopes, A. A survey of semantic web technology for agriculture. *Inf. Process. Agric.* **2019**, *6*, 487–501. [CrossRef]

21. Hernández-Castillo, C.; Guedea-Noriega, H.H.; Rodríguez-García, M.Á.; García-Sánchez, F. Pest recognition using natural language processing. In *Technologies and Innovation. CITI 2019. Communications in Computer and Information Science*; Valencia-García, R., Alcaraz-Mármol, G., Del Cioppo-Morstadt, J., Vera-Lucio, N., Bucaram-Leverone, M., Eds.; Springer: Cham, Switzerland, 2019; Volume 1124, pp. 3–16. [\[CrossRef\]](#)
22. Rodríguez-García, M.Á.; García-Sánchez, F. CropPestO: An ontology model for identifying and managing plant pests and diseases. In *Technologies and Innovation. CITI 2020. Communications in Computer and Information Science*; Valencia-García, R., Alcaraz-Mármol, G., del Cioppo-Morstadt, J., Vera-Lucio, N., Bucaram-Leverone, M., Eds.; Springer: Cham, Switzerland; Guayaquil, Ecuador, 2020; Volume 1309, pp. 18–29. [\[CrossRef\]](#)
23. European Commission. Integrated Pest Management (IPM). Available online: https://ec.europa.eu/food/plant/pesticides/sustainable_use_pesticides/ipm_en (accessed on 29 March 2021).
24. Sankaran, S.; Mishra, A.; Ehsani, R.; Davis, C. A review of advanced techniques for detecting plant diseases. *Comput. Electron. Agric.* **2010**, *72*, 1–13. [\[CrossRef\]](#)
25. Martinelli, F.; Scalenghe, R.; Davino, S.; Panno, S.; Scuderi, G.; Ruisi, P.; Villa, P.; Stroppiana, D.; Boschetti, M.; Goulart, L.R. Advanced methods of plant disease detection. A review. *Agron. Sustain. Dev.* **2015**, *35*, 27–51. [\[CrossRef\]](#)
26. Vishnoi, V.K.; Kumar, K.; Kumar, B. Plant disease detection using computational intelligence and image processing. *J. Plant Dis. Prot.* **2020**. [\[CrossRef\]](#)
27. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent advances in image processing techniques for automated leaf pest and disease recognition—A review. *Inf. Process. Agric.* **2020**. [\[CrossRef\]](#)
28. Nagaraju, M.; Chawla, P. Systematic review of deep learning techniques in plant disease detection. *Int. J. Syst. Assur. Eng. Manag.* **2020**, *11*, 547–560. [\[CrossRef\]](#)
29. Toda, Y.; Okura, F. How convolutional neural networks diagnose plant disease. *Plant Phenomics* **2019**, *2019*, 1–14. [\[CrossRef\]](#)
30. Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* **2019**, *161*, 272–279. [\[CrossRef\]](#)
31. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [\[CrossRef\]](#)
32. Kartikeyan, P.; Shrivastava, G. Review on emerging trends in detection of plant diseases using image processing with machine learning. *Int. J. Comput. Appl.* **2021**, *174*, 39–48. [\[CrossRef\]](#)
33. Liu, J.; Wang, X. Plant diseases and pests detection based on deep learning: A review. *Plant Methods* **2021**, *17*, 22. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Chen, J.-W.; Lin, W.-J.; Cheng, H.-J.; Hung, C.-L.; Lin, C.-Y.; Chen, S.-P. A smartphone-based application for scale pest detection using multiple-object detection methods. *Electronics* **2021**, *10*, 372. [\[CrossRef\]](#)
35. Velásquez, D.; Sánchez, A.; Sarmiento, S.; Toro, M.; Maiza, M.; Sierra, B. A method for detecting coffee leaf rust through wireless sensor networks, remote sensing, and deep learning: Case study of the caturra variety in Colombia. *Appl. Sci.* **2020**, *10*, 697. [\[CrossRef\]](#)
36. Messina, G.; Modica, G. Applications of UAV thermal imagery in precision agriculture: State of the art and future research outlook. *Remote Sens.* **2020**, *12*, 1491. [\[CrossRef\]](#)
37. Zhang, J.; Huang, Y.; Pu, R.; Gonzalez-Moreno, P.; Yuan, L.; Wu, K.; Huang, W. Monitoring plant diseases and pests through remote sensing technology: A review. *Comput. Electron. Agric.* **2019**, *165*, 104943. [\[CrossRef\]](#)
38. Petrellis, N. Plant disease diagnosis for smart phone applications with extensible set of diseases. *Appl. Sci.* **2019**, *9*, 1952. [\[CrossRef\]](#)
39. Cui, S.; Ling, P.; Zhu, H.; Keener, H. Plant pest detection using an artificial nose system: A review. *Sensors* **2018**, *18*, 378. [\[CrossRef\]](#)
40. Hazarika, S.; Choudhury, R.; Montazer, B.; Medhi, S.; Goswami, M.P.; Sarma, U. Detection of Citrus Tristeza virus in mandarin orange using a custom-developed electronic nose system. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9010–9018. [\[CrossRef\]](#)
41. Ruusunen, O.; Jalli, M.; Jauhiainen, L.; Ruusunen, M.; Leiviskä, K. Advanced data analysis as a tool for net blotch density estimation in spring barley. *Agriculture* **2020**, *10*, 179. [\[CrossRef\]](#)
42. Maneesha, A.; Suresh, C.; Kiranmayee, B.V. Prediction of rice plant diseases based on soil and weather conditions. In *Proceedings of the International Conference on Advances in Computer Engineering and Communication Systems. Learning and Analytics in Intelligent Systems*, Greater Noida, India, 19–20 February 2021; Volume 20, pp. 155–165. [\[CrossRef\]](#)
43. Lagos-Ortiz, K.; Medina-Moreira, J.; Paredes-Valverde, M.A.; Espinoza-Morán, W.; Valencia-García, R. An ontology-based decision support system for the diagnosis of plant diseases. *J. Inf. Technol. Res.* **2017**, *10*, 42–55. [\[CrossRef\]](#)
44. Cañadas, J.; del Águila, I.M.; Palma, J. Development of a web tool for action threshold evaluation in table grape pest management. *Precis. Agric.* **2017**, *18*, 974–996. [\[CrossRef\]](#)
45. Toseef, M.; Khan, M.J. An intelligent mobile application for diagnosis of crop diseases in Pakistan using fuzzy inference system. *Comput. Electron. Agric.* **2018**, *153*, 1–11. [\[CrossRef\]](#)
46. Goodridge, W.; Bernard, M.; Jordan, R.; Rampersad, R. Intelligent diagnosis of diseases in plants using a hybrid Multi-Criteria decision making technique. *Comput. Electron. Agric.* **2017**, *133*, 80–87. [\[CrossRef\]](#)
47. Halder, S.; Kumar Singh, S. Knowledge-based expert system for diagnosis of agricultural crops. In *Proceedings of International Conference on Frontiers in Computing and Systems. Advances in Intelligent Systems and Computing*; Springer: Singapore, 2021; Volume 1255, pp. 351–359. [\[CrossRef\]](#)

48. Li, Y.; Luo, Z.; Wang, F.; Wang, Y. Hyperspectral leaf image-based cucumber disease recognition using the extended collaborative representation model. *Sensors* **2020**, *20*, 4045. [\[CrossRef\]](#)
49. Deng, X.; Zhu, Z.; Yang, J.; Zheng, Z.; Huang, Z.; Yin, X.; Wei, S.; Lan, Y. Detection of citrus huanglongbing based on multi-input neural network model of UAV hyperspectral remote sensing. *Remote Sens.* **2020**, *12*, 2678. [\[CrossRef\]](#)
50. Li, K.; Lin, J.; Liu, J.; Zhao, Y. Using deep learning for image-based different degrees of ginkgo leaf disease classification. *Information* **2020**, *11*, 95. [\[CrossRef\]](#)
51. Skawsang, S.; Nagai, M.; Tripathi, N.; Soni, P. Predicting rice pest population occurrence with satellite-derived crop phenology, ground meteorological observation, and machine learning: A case study for the central plain of Thailand. *Appl. Sci.* **2019**, *9*, 4846. [\[CrossRef\]](#)
52. Chen, J.; Liu, Q.; Gao, L. Visual tea leaf disease recognition using a convolutional neural network model. *Symmetry* **2019**, *11*, 343. [\[CrossRef\]](#)
53. Anagnostis, A.; Asiminari, G.; Papageorgiou, E.; Bochtis, D. A convolutional neural networks based method for anthracnose infected walnut tree leaves identification. *Appl. Sci.* **2020**, *10*, 469. [\[CrossRef\]](#)
54. Liang, Q.; Xiang, S.; Hu, Y.; Coppola, G.; Zhang, D.; Sun, W. PD2SE-Net: Computer-assisted plant disease diagnosis and severity estimation network. *Comput. Electron. Agric.* **2019**, *157*, 518–529. [\[CrossRef\]](#)
55. Pantazi, X.E.; Moshou, D.; Tamouridou, A.A. Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. *Comput. Electron. Agric.* **2019**, *156*, 96–104. [\[CrossRef\]](#)
56. Picon, A.; Alvarez-Gila, A.; Seitz, M.; Ortiz-Barredo, A.; Echazarra, J.; Johannes, A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Comput. Electron. Agric.* **2019**, *161*, 280–290. [\[CrossRef\]](#)
57. Arsenovic, M.; Karanovic, M.; Sladojevic, S.; Anderla, A.; Stefanovic, D. Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry* **2019**, *11*, 939. [\[CrossRef\]](#)
58. Sun, R.; Zhang, M.; Yang, K.; Liu, J. Data enhancement for plant disease classification using generated lesions. *Appl. Sci.* **2020**, *10*, 466. [\[CrossRef\]](#)
59. Garcerán-Sáez, J.; García-Sánchez, F. SePeRe: Semantically-enhanced system for pest recognition. In *ICT for Agriculture and Environment. CITAMA2019 2019. Advances in Intelligent Systems and Computing*; Valencia-García, R., Alcaraz-Mármol, G., Cioppo-Morstadt, J., Vera-Lucio, N., Bucaram-Leverone, M., Eds.; Springer: Cham, Switzerland, 2019; Volume 901, pp. 3–11. [\[CrossRef\]](#)
60. Labaña, F.M.; Ruiz, A.; García-Sánchez, F. PestDetect: Pest recognition using convolutional neural network. In *ICT for Agriculture and Environment. CITAMA2019 2019. Advances in Intelligent Systems and Computing*; Valencia-García, R., Alcaraz-Mármol, G., Cioppo-Morstadt, J., Vera-Lucio, N., Bucaram-Leverone, M., Eds.; Springer: Cham, Switzerland, 2019; Volume 901, pp. 99–108. [\[CrossRef\]](#)
61. Colombo-Mendoza, L.O.; Valencia-García, R.; Colomo-Palacios, R.; Alor-Hernández, G. A knowledge-based multi-criteria collaborative filtering approach for discovering services in mobile cloud computing platforms. *J. Intell. Inf. Syst.* **2020**, *54*, 179–203. [\[CrossRef\]](#)
62. Rodríguez-García, M.Á.; Valencia-García, R.; Colomo-Palacios, R.; Gómez-Berbís, J.M. BlindDate recommender: A context-aware ontology-based dating recommendation platform. *J. Inf. Sci.* **2019**, *45*, 573–591. [\[CrossRef\]](#)
63. Rodríguez Iglesias, A.; Egaña Aranguren, M.; Rodríguez González, A.; Wilkinson, M.D. Plant-Pathogen Interactions Ontology (PPIO). In *Proceedings of the International Work-Conference on Bioinformatics and Biomedical Engineering IWBBIO 2013*, Granada, Spain, 18–20 March 2013; Rojas, I., Ortuño Guzman, F.M., Eds.; Copicentro Editorial: Granada, Spain, 2013; pp. 695–702.
64. Walls, R.; Smith, B.; Elser, J.; Goldfain, A.; Stevenson, D.W.; Jaiswal, P. A plant disease extension of the Infectious Disease Ontology. In *Proceedings of the 3rd International Conference on Biomedical Ontology (ICBO 2012)*, KR-MED Series, Graz, Austria, 21–25 July 2012; Cornet, R., Stevens, R., Eds.; CEUR-WS.org: Graz, Austria, 2012; pp. 1–5.
65. Dalvi, P.; Mandave, V.; Gothkhindi, M.; Patil, A.; Kadam, S.; Pawar, S.S. Overview of agriculture domain ontologies. *Int. J. Recent Adv. Eng. Technol.* **2016**, *4*, 5–9.
66. Caracciolo, C.; Stellato, A.; Morshed, A.; Johannsen, G.; Rajbhandari, S.; Jaques, Y.; Keizer, J. The AGROVOC linked dataset. *Semant. Web.* **2013**, *4*, 341–348. [\[CrossRef\]](#)
67. Beck, H.W.; Kim, S.; Hagan, D. A crop-pest ontology for extension publications. In *Proceedings of the 2005 EFITA/WCCA Joint Congress on IT in Agriculture*, Vila Real, Portugal, 25–28 July 2005; pp. 1169–1176.
68. Jonquet, C.; Toulet, A.; Arnaud, E.; Aubin, S.; Dzalé Yeumo, E.; Emonet, V.; Graybeal, J.; Laporte, M.-A.; Musen, M.A.; Pesce, V.; et al. AgroPortal: A vocabulary and ontology repository for agronomy. *Comput. Electron. Agric.* **2018**, *144*, 126–143. [\[CrossRef\]](#)
69. Xiaoxue, L.; Xuesong, B.; Longhe, W.; Bingyuan, R.; Shuhan, L.; Lin, L. Review and trend analysis of knowledge graphs for crop pest and diseases. *IEEE Access* **2019**, *7*, 62251–62264. [\[CrossRef\]](#)
70. Lacasta, J.; Lopez-Pellicer, F.J.; Espejo-García, B.; Nogueras-Iso, J.; Zarazaga-Soria, F.J. Agricultural recommendation system for crop protection. *Comput. Electron. Agric.* **2018**, *152*, 82–89. [\[CrossRef\]](#)
71. Titiya, M.D.; Shah, V.A. Ontology based expert system for pests and disease management of cotton crop in India. *Int. J. Web. Portals* **2018**, *10*, 32–49. [\[CrossRef\]](#)
72. Jearanaiwongkul, W.; Anutariya, C.; Andres, F. An ontology-based approach to plant disease identification system. In *Proceedings of the Proceedings of the 10th International Conference on Advances in Information Technology—IAIT 2018*, Bangkok, Thailand, 10–13 December 2018; ACM Press: New York, NY, USA, 2018; pp. 1–8. [\[CrossRef\]](#)

73. Somodevilla, M.J.; Vilariño Ayala, D.; Pineda, I. An overview on ontology learning tasks. *Comput. Sist.* **2018**, *22*, 137–146. [CrossRef]
74. Khadir, A.C.; Aliane, H.; Guessoum, A. Ontology learning: Grand tour and challenges. *Comput. Sci. Rev.* **2021**, *39*, 100339. [CrossRef]
75. Xiang, Z.; Zheng, J.; Lin, Y.; He, Y. Ontorat: Automatic generation of new ontology terms, annotations, and axioms based on ontology design patterns. *J. Biomed. Semant.* **2015**, *6*, 4. [CrossRef]
76. Laaz, N.; Wakil, K.; Gotti, S.; Gotti, Z.; Mbarki, S. An automatic generation of domain ontologies based on an MDA approach to support big data analytics. In *Advancements in Model-Driven Architecture in Software Engineering*; IGI Global: Hershey, PA, USA, 2021; pp. 26–45. [CrossRef]
77. Lubani, M.; Noah, S.A.M.; Mahmud, R. Ontology population: Approaches and design aspects. *J. Inf. Sci.* **2019**, *45*, 502–515. [CrossRef]
78. Petasis, G.; Karkaletsis, V.; Paliouras, G.; Krithara, A.; Zavitsanos, E. Ontology population and enrichment: State of the art. In *Proceedings of the Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 134–166.
79. Qiu, J.; Chai, Y.; Liu, Y.; Gu, Z.; Li, S.; Tian, Z. Automatic non-taxonomic relation extraction from big data in smart city. *IEEE Access* **2018**, *6*, 74854–74864. [CrossRef]
80. Reynaud, J.; Toussaint, Y.; Napoli, A. Redescription mining for learning definitions and disjointness axioms in linked open data. In *Proceedings of the International Conference on Conceptual Structures*, Marburg, Germany, 1–4 July 2019; Springer: Cham, Switzerland; Marburg, Germany, 2019; pp. 175–189. [CrossRef]
81. Nguyen, T.H.; Tettamanzi, A.G.B. Grammatical evolution to mine OWL disjointness axioms involving complex concept expressions. In *Proceedings of the 2020 IEEE Congress on Evolutionary Computation (CEC)*, Glasgow, UK, 19–24 July 2020; IEEE: Glasgow, UK, 2020; pp. 1–8. [CrossRef]
82. Ochoa, J.L.; Valencia-García, R.; Perez-Soltero, A.; Barceló-Valenzuela, M. A semantic role labelling-based framework for learning ontologies from Spanish documents. *Expert Syst. Appl.* **2013**, *40*, 2058–2068. [CrossRef]
83. Sad-Houari, N.; Taghezout, N.; Nador, A. A knowledge-based model for managing the ontology evolution: Case study of maintenance in SONATRACH. *J. Inf. Sci.* **2019**, *45*, 529–553. [CrossRef]
84. Abdou, M.; AbdelGaber, S.; Farhan, M. A semi-automated framework for semantically annotating web content. *Futur. Gener. Comput. Syst.* **2018**, *81*, 94–102. [CrossRef]
85. Zhang, C.; Chen, J.; Zhang, L.; Chen, S.; Zhang, Z. A new semantic annotation approach for software vulnerability source code. *Int. J. Simul. Process Model.* **2021**, *16*, 1–13. [CrossRef]
86. Rodríguez-García, M.Á.; Valencia-García, R.; García-Sánchez, F.; Samper-Zapater, J.J. Creating a semantically-enhanced cloud services environment through ontology evolution. *Futur. Gener. Comput. Syst.* **2014**, *32*, 295–306. [CrossRef]
87. Rodríguez-García, M.Á.; Valencia-García, R.; García-Sánchez, F.; Samper-Zapater, J.J. Ontology-based annotation and retrieval of services in the cloud. *Knowl.-Based Syst.* **2014**, *56*, 15–25. [CrossRef]
88. Albukhitan, S.; Helmy, T.; Alnazer, A. Arabic ontology learning using deep learning. In *Proceedings of the Proceedings of the International Conference on Web Intelligence—WI '17*; ACM Press: New York, NY, USA, 2017; pp. 1138–1142. [CrossRef]
89. Noy, N.F.; McGuinness, D.L. Ontology Development 101: A Guide to Creating Your First Ontology. Available online: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf (accessed on 6 March 2021).
90. Food and Agriculture Organization (FAO). AGROVOC. Available online: <http://aims.fao.org/vest-registry/vocabularies/agrovoc> (accessed on 6 March 2021).
91. W3C OWL Working Group. OWL 2 Web Ontology Language Document Overview (Second Edition). Available online: <https://www.w3.org/TR/owl2-overview/> (accessed on 6 March 2021).
92. Ministerio de Agricultura Alimentación y Medio Ambiente. Metodología de la Estadística Sobre Superficies y Producciones Anuales de Cultivos. Available online: https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/NotasMetodológicasSuperficiesyproduccionesanualesdecultivos_tcm30-122273 (accessed on 6 March 2021).
93. Ministerio de Medio Ambiente y Medio Rural y Marino. *Patógenos de Plantas Descritos en España*, 2nd ed.; Sociedad Española de Fitopatología: Madrid, Spain, 2010; ISBN 978-84-491-0954-6.
94. Salton, G.; McGill, M.J. *Introduction to Modern INFORMATION Retrieval*; McGraw-Hill, Inc.: New York, NY, USA, 1983; ISBN 0070544840.
95. Goldstein, A.; Fink, L.; Ravid, G. A framework for evaluating agricultural ontologies. *arXiv* **2019**, arXiv:1906.10450.
96. Loizou, E.; Karelakis, C.; Galanopoulos, K.; Mattas, K. The role of agriculture as a development tool for a regional economy. *Agric. Syst.* **2019**, *173*, 482–490. [CrossRef]
97. Shennan, C.; Krupnik, T.J.; Baird, G.; Cohen, H.; Forbush, K.; Lovell, R.J.; Olimpi, E.M. Organic and conventional agriculture: A useful framing? *Annu. Rev. Environ. Resour.* **2017**, *42*, 317–346. [CrossRef]