Improving access to research outcomes for innovation in agriculture and forestry: the VALERIE project

Luca Bechini,[1] Nicole Koenderink,[2] Hein F.M. ten Berge,[3] Wim Corre,[3] Frits K. van Evert,[3] Arianna Facchi,[1] Olfa Gharsallah,[1] Elena Gorriz-Mifsud,[4] Carlo Grignani,[5] Michael den Herder,[6] Yolaine Hily,[7] Eric Justes,[7] Aurélien Lepennetier,[8] Barbara Moretti,[5] Paul Newell-Price,[9] Luca Nonini,[1] Roberto Oberti,[1] Sonia Ramonteu,[8] Mercedes Rois,[4] Frank de Ruijter,[3] Dario Sacco,[5] Peter M. Schuler,[10] Don Willems,[2] Anneke Zandstra,[2] Jan Top[2]

[1]Department of Agricultural and Environmental Sciences, University of Milan, Milan, Italy; [2]Wageningen Food & Biobased Research, Wageningen; [3]Wageningen Plant Sciences Group, Wageningen, The Netherlands; [4]European Forest Institute, Barcelona, Spain; [5]Department of Agricultural, Forestry, and Food Sciences, University of Turin, Turin, Italy; [6]European Forest Institute, Joensuu, Finland; [7]Unité Mixte de Recherche Agroécologie, Innovations, Territoire, Institut National de la Recherche Agronomique, Castanet Tolosan; [8]Association de Coordination Technique Agricole, Paris, France; [9]ADAS Gleadthorpe, Meden Vale, Mansfield, UK; [10]it-objects GmbH, Hagen, Germany

Abstract

Many excellent results are obtained in agricultural and forestry research projects, but their practical adoption is often

limited. The aim of the European project VALERIE is to increase the transfer and application of innovations produced by research in agriculture and forestry, by facilitating their integration into management practices. The project is still ongoing and the results illustrated in this paper are still temporary and subject to being improved. Here we present the methodology used in VALERIE to extract and summarise knowledge for innovation from research documents with the aim of making it available to final users through *ask-Valerie.eu*; we also report on current progress. The tasks associated with extracting and summarising knowledge are centred on: i) an ontology; ii) a document base; and iii) a system (*ask-Valerie.eu*) that allows users to effectively search the document base. An ontology defines a set of concepts and the relations between them. The VALERIE ontology is built by experts in the agricultural and forestry domain and contains 6169 concepts (21st October 2016). The document base is the collection of documents in which the system searches. The VALERIE document base includes scientific and practical documents derived from various sources, written in any of a number of languages. All documents contained in the document base are annotated using the ontology: each term (a word or a short phrase) in the document that matches a concept in the VALERIE-ontology is linked to that concept. Annotation is an automated process that takes place whenever a document is added to the document base. The document base contains 4278 documents (October 2016). Among them, there are 201 mini-factsheets written by members of the VALERIE project, each describing an innovation with: a short description of the innovation, a list of correlated projects, and some links to scientific and practical documents. *ask-Valerie.eu* searches documents and fragments of text from the document base that address the user's query. *ask-Valerie.eu* mimics the dialogue between a practitioner and an expert and achieves this functionality by: i) supporting the practitioner in articulating the question (it completes terms that the user starts to type and suggests other possibly relevant terms); ii) expanding the query using synonyms; iii) extracting and ranking text fragments from the documents.

Introduction

There is a compelling need for agricultural and forestry research to play a significant role in the future to meet the challenges of increased demand for food and fibre. This should be balanced against the need to deliver other vital services provided by land management, such as maintaining valued habitats. If this role is to be fulfilled, innovative research outputs need to be made accessible to end-users so that they can be applied in practice. Many EU and nationally funded research projects in the fields of agriculture and forestry provide excellent scientific results. However, outreach and interpretation of these results into *ready-to-use* farming and forestry formats is limited. Among the causes of limited dissemination of research outputs are the lack of a structured catalogue of research products and a limited diffusion in trade journals or in the form of practically-oriented fact sheets devoted to final users such as farmers, advisors, and foresters.

More broadly, three major developments have led to an increased need for effective uptake of research results into farming practice. Firstly, many threats facing agriculture, such as climate change, the vulnerability of food systems, food and water insecurity, soil degradation and the loss of biodiversity, represent large-scale challenges for the industry. Secondly, the policy agenda in Europe has changed with reorientation of the CAP (Common Agricultural Policy) and the increasing emphasis on the economic, environmental and social dimensions of sustainability. Thirdly, extension policy reforms of the 1990s in many EU countries have seen a move away from government-driven and funded Agricultural Knowledge and Innovation Systems (AKIS) towards pluralistic, market-driven, multi-actor systems, in which private actors have come to play a larger role (Rivera and Sulaiman, 2009). As a result, there are increasing demands on farmers to produce more with less, with *more* often explicitly including public goods and services, and they have to meet these demands in an environment where there is an increasing disconnection between research and farming.

The main aim of the European project VALERIE (2014-2017; http://www.valerie.eu) is to improve the accessibility to and availability of new knowledge for innovation in agriculture and forestry produced by European and national research, and to facilitate its integration into farming and forestry practice. More specifically, we are developing a digital *assistant expert* in the field of agriculture and forestry (*ask-Valerie.eu*) that could help farmers and foresters to find information on innovation. In particular, the digital expert will allow users to articulate queries, find all relevant results (documents/fragments of documents, projects, people, equipment, software), and rank them. Appendix Table 1 gives a short description of a potential case of *ask-Valerie.eu* being used by organic dairy farmers in Wales. Knowing that the old linear model of innovation and adoption (from scientists to users through extension services) proved to be unsuccessful in addressing complex agri-environmental issues and has been progressively replaced by an interactive model of networking systems (Leeuwis, 2004; Knickel *et al.*, 2009; Smits *et al.*, 2010), the VALERIE project also runs case studies that include a social science perspective. This manuscript does not describe these activities, but focuses on describing the approach to summarising and extracting knowledge on innovations from existing literature. Summarising and extracting knowledge are a *conditio sine qua non* for successive actions, meaning that using *ask-Valerie.eu* is only one action in a longer sequence of events in knowledge sharing. These events might include: check if the innovation fits with actual farm organisation, including labelling schemes imposed by retailers; find which partners are needed to implement the innovation (*e.g.* find who sells the innovation; find who offers services to apply it); consult an advisor to obtain additional information about the innovation, or take a course; find if there are financial incentives to support the innovation; connect with other farmers who have used or are starting to use the innovation; after the innovation was applied, verify its results, both internally and externally (with clients, and with other referents such as advisors and input suppliers); if needed, modify the innovation to meet local requirements. Our point of view in developing *ask-Valerie.eu* is that, if users do not know what kind of knowledge is available, it will be impossible for them to make use of it.

The objective of this article is to present the VALERIE methodology for extracting, summarising and sharing knowledge for innovation in agriculture and forestry, and to report current progress in these ongoing activities. We underline that the state of *ask-Valerie.eu* and its components as described refers to late 2016, when a limited document base is available for development and testing purposes only, and the domain ontology is still being refined and expanded.

Materials and methods

The VALERIE project

To improve the accessibility and availability of new knowledge for innovation in agriculture and forestry, the VALERIE project combines a science-driven approach and a stakeholder-driven approach, facilitating knowledge exchange between practitioners (farmers, foresters, and advisors) and researchers to improve the relevance and uptake of research outputs.

In the stakeholder-driven approach, we work with practitioners in ten case studies to identify knowledge gaps, to assess technical and economic viability of innovative solutions, and to reveal barriers to uptake. Case studies are organised around a particular supply chain, a farming or forestry sector, or a landscape, and so cover different scales and dimensions.

The focus of this article is on the science-driven approach, in which we extract, review, and summarise knowledge from ongoing and completed national, international, and European research projects, and develop the digital expert *ask-Valerie.eu* to improve access to the extracted and summarised knowledge. *ask-Valerie.eu* will not only make new knowledge accessible to the end-users, but will also enable them to share their empirical knowledge, experience and views with peers across Europe. The science-driven work is organised around six themes: i) Crop rotation, soil cover management and integrated pest management; ii) Ecosystem and social services in agriculture and forestry; iii) Soil management as an integrated agro-ecological system; iv) Water management in agriculture and forestry; v) Integrated supply chain services and tools, innovative farm management; vi) Recycling and smart use of biomass and food waste, in particular waste generated during primary production.

The tasks of extracting and summarising knowledge in VALERIE are centred on three main components: i) an ontology (Gruber, 2009); ii) a document base consisting of annotated and non-annotated documents; and iii) a system (*ask-Valerie.eu*) that allows users to search the document base effectively.

A short overview of how these three components are used is given here, with more details provided in the next paragraphs. The

ontology is a structured collection of concepts used in agriculture and forestry that covers the entire domain explored in VALERIE. The ontology is used to *annotate* documents contained in the document base that consists of scientific papers, articles in trade journals, practical fact sheets, and manuals. In the annotation process, terms found in documents are matched with terms found in the ontology. As a result, each document is collocated within the VALERIE domain. Once the documents are annotated, they can be found by *ask-Valerie.eu* users through queries based on concepts present in the ontology. As will be shown below, one of the most important advantages of *ask-Valerie.eu* is that the user can navigate the results by mimicking the interaction with a real expert.

Ontology

An ontology defines a set of representational primitives (concepts) and the relations between these primitives (Gruber, 2009). The VALERIE domain consists of the above-mentioned six themes within the agronomy and forestry sectors. The VALERIE domain ontology is defined in the SKOS-format and consists of concepts (terms) from these themes and this domain. Each concept may have synonyms (skos:altLabel) in multiple languages. The concepts have been interrelated in two ways. First, a hierarchical organisation has been added to the terms, using an *is a kind of* relation (skos:narrower; http://www.w3.org/TR/skos-reference/). Concepts may have more than one parent. Second, the relation (skos:related) has been added to allow for other, more generic connections between concepts.

The VALERIE domain ontology has been created using the ROC+ method as described in Willems *et al*. (2015). One of the key benefits of the ROC+ method is that the VALERIE scientists can optimally draw concepts and partial concept trees from already existing ontologies such as Agrovoc (http://aims.fao.org/standards/agrovoc/linked-open-data) and Eurovoc (http://eurovoc.europa.eu/drupal/). This way, work already carried out to model relevant domain ontologies is reused in the VALERIE-ontology.

Document base

Types of documents belonging to VALERIE's document base

The *ask-Valerie.eu* document base is the collection of documents in which *ask-Valerie.eu* searches. From this base, searches by *ask-Valerie.eu* will return one or more documents, one or more text fragments, or a combination of these. It is important to underline that *ask-Valerie.eu* does not store copies of the documents. Documents are regarded as part of the *ask-Valerie.eu* document base once they have been annotated. Annotation results in a stored index (set of terms that refer to concepts in the ontology). The original documents remain in their original repositories.

VALERIE's document base includes both scientific and practical documents derived from various sources, and in multiple languages (Italian, English, French, Finnish, Spanish, and Dutch). In contrast to standard search engines, VALERIE does not use the entire web as its source of information. Rather, VALERIE experts select sources containing documents related to agriculture and forestry that meet certain usefulness and relevance criteria. So far, we have identified three categories of documents of increasing quality.

First, documents from selected repositories. These repositories include collections of scientific articles published in agricultural and forestry journals; collections of articles published in trade journals; repositories of practical fact sheets; and repositories of research project outcomes, like CORDIS.

Second, documents individually considered relevant by experts. To prepare mini-factsheets, among the many documents contained in existing repositories, VALERIE experts select some that are particularly valuable, for example scientific reviews, or high-quality practical documents. Compared to the bulk of documents contained in repositories, these manually selected documents are of higher importance because they are selected for their relevance and scientific accuracy by experts.

Third, mini-factsheets prepared by VALERIE. For precisely identified innovations, VALERIE experts prepare short factsheets (see paragraph Innovations and mini-factsheets for more details). These are the documents of highest relevance in the document base in the sense that they are assumed to give the most specific answers to users' questions.

Annotation

All documents contained in the document base are automatically annotated using VALERIE's domain ontology in the annotation tool AnnA (Willems *et al*., 2015). During annotation, each term in the document that matches a concept in the VALERIE-ontology is tagged. As a result, each document contains links to the ontology. If a term in a document is not present as a concept in the ontology, it will not be annotated. The annotation of a text represents the computer's understanding (model) of that text. This model allows the *virtual assistant* to communicate with the user because it knows the relations between terms. Purely text-based search engines do not have this intelligence.

In addition to knowing the domain language (using concepts such as *soil*, *potato* or *tillage*), the virtual assistant should also be aware of the role that text fragments or other pieces of information play in the query-and-answer process between a farmer or forester and the advisor. The most basic roles in this conversation are *question* and *answer*. Since we are focusing on knowledge for innovation, we have labelled these as respectively *innovation challenges* and *innovations*. If in a document a piece of text is identified as the description of an *innovation challenge* and if it uses terms that also occur in the question asked, the associated *innovation* text is probably the most relevant and valuable answer that can be found automatically. In the ontology, we have specified cause-effect relations that will be used by the system to automatically identify parts of the text documents that describe innovations and innovation challenges.

Innovations and mini-factsheets

Our solution to identify documents (and fragments of text within documents) dealing with innovations is to prepare a list of innovations in agriculture and forestry. The idea is that, once innovations have been properly defined, we will take advantage of the VALERIE ontology to identify more documents in the document base describing these innovations.

Many different definitions of innovation can be given. Within the VALERIE project, an innovation is *a practice, a solution or a tool that can be implemented to address a specific problem for a farmer, a forester, or an advisor*. A given practice or solution or tool might be innovative in one place and not in another. For example, no-tillage techniques are widely adopted in many contexts, but may still be innovative in certain regions or with farmers inexperienced with the approach.

Our list of innovations was compiled separately by theme. To do this, we have consulted experts in the six themes. The main criterion in inviting experts was to ensure a good coverage, both geographically (experts from different European countries) and scientifically (different topics within the theme). Currently we have

involved 34 scientists from eight countries and seventeen institutions. Eleven experts have a role in VALERIE, while 23 do not belong to the project.
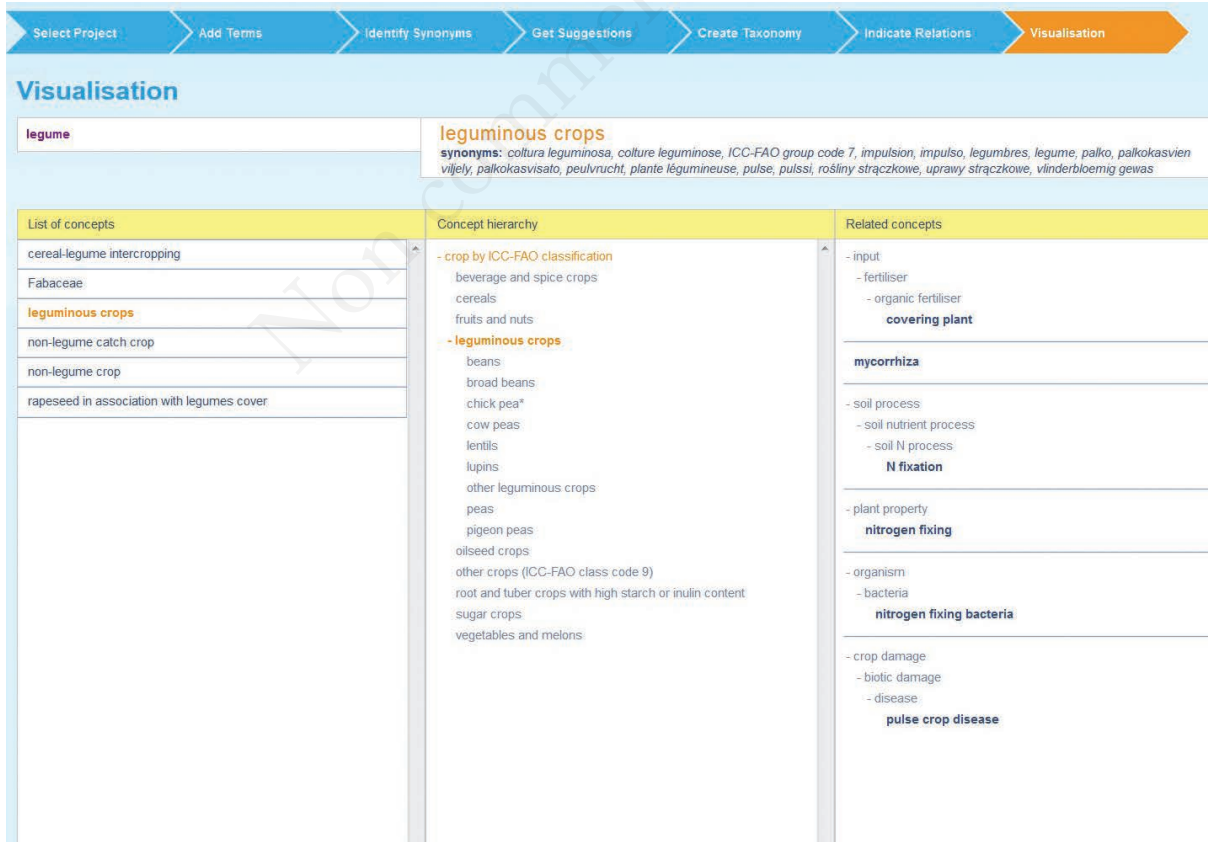
For each innovation, we aimed to compile a simplified factsheet; namely a VALERIE *mini-factsheet*. A mini-factsheet is a short document that presents an overview of the innovation. It links to i) documents describing the innovation, ii) projects where the innovation was studied or developed, iii) the process or problem the innovation aims to address (the *innovation challenges*; *e.g.* weed control, efficient water use), and the iv) related concepts (terms hand-picked by experts from the ontology to facilitate searching with *ask-Valerie.eu*). Mini-factsheets are a compromise between (project) resource availability and the aim to cover a wide range of thematic topics.

We use several criteria in choosing documents to be referred in mini-factsheets. The first is to include links to both scientific and practical documents (3-4 documents per type). These two types have different functions. While scientific papers explain the science behind the innovation, or may confirm the innovation's effectiveness, practical documents often instruct on how, when and where to use the innovation. Another criterion is to avoid documents that provide a general context; we prefer documents that are focused specifically on innovations. Third, we prefer documents that have a large experimental basis, avoiding those that describe a particular situation. For example, a document that refers to multisite/multi-year experiments is preferred to one that is based only on one year and one site.

ask-Valerie.eu

A principal aim of the VALERIE project is to develop *ask-Valerie.eu,* a digital assistant expert, combining the processing and storage power of computer systems with the knowledge of domain experts. *ask-Valerie.eu* mimics the dialogue between practitioner and expert (Willems *et al.*, 2015). Technically, *ask-Valerie.eu* achieves this by executing a number of processes. Firstly, it supports the practitioner in articulating the user question. It completes terms that the user starts to type; suggests other possibly relevant terms; automatically detects synonyms or more specific terms in text fragments; and extracts and ranks text fragments (Willems *et al.*, 2015).

The *ask-Valerie.eu* system consists of a front-end that facilitates communication with the user and a back-end that processes the questions posed, generating the answers by consulting the document base and the ontology. Since we intend to mimic a natural conversation between the system and the user, we allow the user to simply state a question as a full sentence, as if he/she is speaking to a human expert. The system then identifies all terms in this question that it knows from the ontology. This is done continuously by suggesting ontological terms while the user is typing, but also by suggesting other possibly relevant terms. This facilitates the user in gradually refining his or her question. In addition, *ask-Valerie.eu* selects text fragments (*snippets*) from the document base and ranks these according to the expected relevance for the user. The ranking of text fragments is based on the occurrence of concepts from the query. These fragments may support the user in refining the question, in order to obtain best-tailored results.

Figure 1. Screenshot of ROC+ showing part of the ontology dealing with pulse crops.

Results

Ontology

The creation of an ontology is an iterative process, which will be developed with at least four versions of the VALERIE domain ontology foreseen during the four year project. Currently, version 7 of the VALERIE ontology is available online as Linked Open Data at http://www.foodvoc.org (consulted on 21st October 2016), the next version of the VALERIE ontology is under construction. The latest version of the ontology consists of 6169 concepts, 30,515 synonyms and translations, 6194 terms connected via a skos:narrower relation and 3090 terms connected via a skos:related relation. All concepts in the ontology are translated into French, Italian, Finnish, Spanish, Polish and Dutch, and frequently synonyms in these languages have been added. Thus *ask-Valerie.eu* is a multi-lingual platform, where each user can search for answers in their own languages, and find answers in preselected languages. That way, an Italian farmer posing questions in Italian, may find results in French, Spanish and English if these are among his/her *usable* languages, or even in Polish, Finnish and Dutch.

A small part of the ontology, shown as an example, is depicted in Figure 1 using a screenshot of the ROC+ software.

Document base

Documents collected so far

To enable developing and testing of the *ask-Valerie.eu* system, we have built a temporary, limited document repository composed of expert-selected documents and VALERIE mini-factsheets. The number of documents is currently (July 2016) 4278: 201 mini-factsheets; 92 books; 40 book chapters; 144 conference papers; 2310 journal articles; 25 magazine articles; 36 presentations; 129 reports; 20 theses; 7 videos; 357 web pages; and 917 other documents, including newspaper articles, blog posts, and unclassified attachments. This material is regarded sufficient for use in mimicking searches by *ask-Valerie.eu*, as well as to address the specific agronomic/forestry issues brought forward by the project case studies.

Annotation

Figure 2 shows the result of an example of automatic annotation carried out using VALERIE's ontology. The terms with a grey background (*irrigation*, *farmer*, *irrigation scheduling*, *identification*, *farm level*, *water shortage*, *timing*, *water application*, *reduction*, *tolerance level*) are present in the ontology and for this reason were tagged. This means that a link is established between the document and these concepts of the ontology (this view is not presented to *ask-Valerie.eu* users).

Innovations and mini-factsheets

The *list of innovations* contains 479 innovations, categorised into six themes. A hierarchical classification of innovations is not strictly needed for the functioning of *ask-Valerie.eu*, because it is the ontology that guides users in finding innovations. However, to communicate project results outside the *ask-Valerie.eu* platform (for example with expert panels, and in this manuscript), we decided to assign a sub-theme to each innovation.

Table 1 shows the number of innovations by theme and sub-theme, with some examples. More than half of these innovations are already connected to at least one project (mostly European projects FP5, FP6 or FP7).

Theme 1 is dedicated to *Crop rotations including soil cover management and Integrated Pest Management*. Given that crop rotations and soil cover management are key levers in integrated pest management strategies, this theme focuses on pest management (animal pests, pathogenic microorganisms, and weeds). Sub-themes have been defined according to the components of integrated plant protection strategies: i) pest detection and monitoring, including the determination of pest pressure and risk assessment; ii) models and decision support systems that can underpin systematic management strategies; and iii) existing pest management methods classification (*i.e.* grouping the different management strategies according to their mode of action, knowing that there is

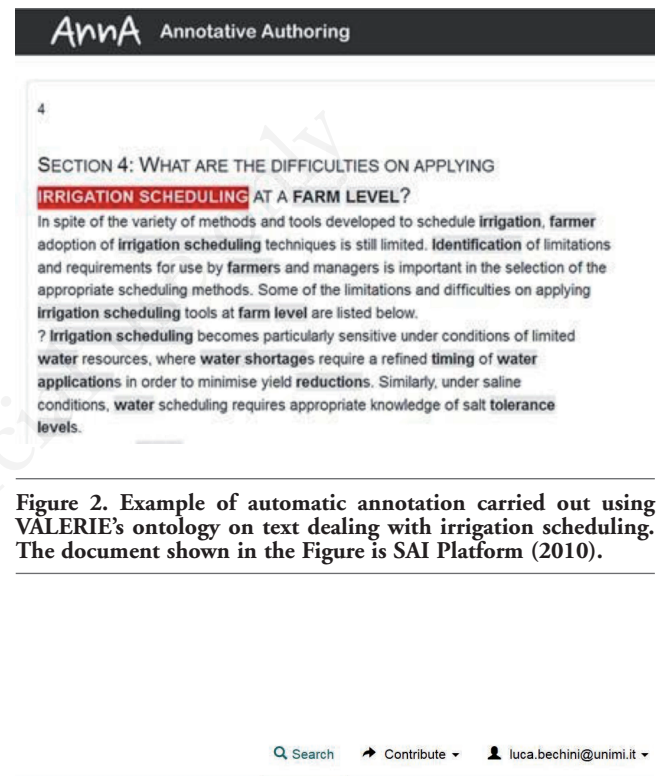**Figure 2. Example of automatic annotation carried out using VALERIE's ontology on text dealing with irrigation scheduling. The document shown in the Figure is SAI Platform (2010).**

Figure 3. An *ask-Valerie.eu* screenshot showing autocompletion: while the user types the query, the system suggests terms that are concepts in the ontology. The system accepts queries in selected languages.

Table 1. Number of innovations identified to date (October 2016) in the VALERIE project, distinguished by theme and sub-theme.

Theme	Sub-theme	Innovations (n)	Examples
Crop rotation, soil cover management and integrated pest management	Biological control	19	Trap plants to control cyst nematodes
	Chemical control	26	Variable rate crop protection treatments
	Cultural control	24	Install buffer strips inside the field
	Genetic control	6	Herbicide-tolerant varieties
	Integrated pest management	16	Integrated management of Tobacco Rattle Virus in potato production
	Management of agroforestry systems	1	Intercropping olives with other economically viable crops in innovative designs to enhance economic sustainability
	Pest monitoring and detection	11	Disease detection by imaging from unmanned aerial vehicle
	Physical control	14	Camera-steered weeding machines
Ecosystem and social services in agriculture and forestry		4	Evaluation of forest ecosystem services
Soil management as an integrated agro-ecological system	Breeding management	10	Regional management of animal waste through exchange of animal manure to equilibrate soil fertility at a regional scale
	Fertilisation management	57	Use of nanoparticles fertilisers to stimulate growth and yield of plants
	Fertility assessment	9	Combined electric conductivity, organic matter and pH measurement for soil mapping
	Precision agriculture	15	Aerial imagery to improve nitrogen fertilisation management
	Soil cover management	17	Use of biodegradable mulching
	Soil tillage	11	Direct sowing into living mulch/permanent cover
Water management in agriculture and forestry	Agricultural drainage	6	Software models for the design, installation and management of drainage networks
	Crop management to reduce water consumption	2	Crop management to decrease irrigation water consumption
	Ecosystem services connected to the farm irrigation system	2	Integration of fish and shrimp farming into harvestable crops and irrigation networks
	Forest water management	9	Mapping water use efficiency of forests to assess drought effects
	Irrigation system materials, machinery and management	46	Irrigation scheduling based on matric soil water potential sensors
	Irrigation water harvesting	6	Biological covers to reduce evaporation losses from on-farm reservoirs
	Protection of water bodies from agricultural pollution	1	Best farm practices to protect water bodies from agricultural pollution
	Soil management to decrease soil evaporation/increase soil water infiltration	4	Use of biopolymers to decrease soil evaporation
	Use of non-conventional water in agriculture	10	On-farm compact wastewater treatment systems
Integrated supply chain services and tools, innovative farm management	Agroforestry supply chain	1	Domestication and development of baobab and tamarind
	Animal husbandry	3	Grazing management to reduce forest fires risk
	Biomass production	2	Use of pruning residues as a source of woody biomass
	Certification	3	Improved methods for branding agroforestry-derived products
	Food supply chain	8	Innovation for urban food planning: short chain delivery of food for urban-peri-urban areas
	Forest management	12	Using light detection and ranging remote sensing to measure aboveground biomass in tree biomes
	Forest wood chain	7	Sustainable use of wood-based biomass
	Management of agroforestry systems	13	Short rotation coppice integrated in an alley cropping systems as a source of woody biomass
	Models and decision support systems	7	Decision support tools for forest risk management
	Natural risks and risk management	4	Forest management to adapt to climate change impact
	Non-wood forest products	5	Production of birch sap juice
	Organic farming and low-input systems	2	Copper-free low input and organic farming systems
	Post-harvest technology	1	Aladin simulation model: quantifying advantages of shelf-life extension like packaging, cooling
	Precision farming	1	Harvesting precision information for better farming
	Quality control and management	3	A monitoring system to prevent virus contamination of food
	Social innovations	1	Sustainable agriculture in Alpine regions
	Supply chain management	1	Cascade use of wood/biomass
	Wood processing	2	Improve the quality of sawn softwood
Recycling and smart use of biomass and food waste, in particular waste generated during primary production		53	Ammonia stripping from manure

no unique classification and that some strategies cannot be strictly classified because of their complexity).

The boundaries of Theme 2 (*Ecosystem and social services in agriculture and forestry*) are hard to define, as are the innovations within this theme, as ecosystem services include any good or service that is provided within the agriculture and forestry domains, from food and timber to flood control and recreation. Innovative concepts in this theme are often relevant to policy makers, but are less applicable to practitioners in the field unless linked to subsidies for sustainable agriculture or other payments for ecosystem services. The boundaries of this theme are also confused by the fact that many of the positive agro-environmental effects of management practices that can be considered ecosystem services are covered by other themes (*e.g.* soil management). For this reason, the number of innovations collected in this theme is still relatively small.

Theme 3 (*Soil management as an integrated agro-ecological system*) includes innovations acting on the soil that can have a direct effect on crop growth or on the environmental impact of agricultural and silvo-pastoral activities. Most of the technologies included in the sub-theme *Nutrient management* relate to the availability of new fertilisers or to new fertiliser application techniques. *Breeding management*, *Soil cover management* and *Soil tillage* represent some well-established principles for which most of the innovations include the development of new machinery or management practices. Finally, *Precision agriculture* and *Fertility assessment* represent technologies based on Global Positioning Systems (GPS); monitoring of soil and weather; and remote and proximal sensing (using soil and crop reflectance) have been made available in the last few years to farmers and third party service providers.

Theme 4 is devoted to the thematic domain of *Water management in agriculture and forestry*. Many methods are available to increase water use efficiency in agriculture, which is often very low in many parts of Europe. These methods range from the application of advanced irrigation technologies and the adoption of innovative irrigation scheduling systems, to the management of crops to decrease their water consumption or the adoption of drought-tolerant cultivars. All these innovations are included in the sub-theme *Irrigation system materials, machinery and management*. Moreover, water use efficiency can be also increased by using alternative sources of water, as illustrated by innovations included within the sub-theme *Use of non-conventional water in agriculture*, or by improving water harvesting through innovations reported in the sub-theme *Irrigation water harvesting*, as well as through careful soil management, as highlighted by the sub-Theme *Soil management to decrease soil evaporation/increase soil water infiltration*. In some regions, drainage is more of an issue than water availability: innovations in this field are listed in the *Agricultural drainage* sub-theme. Water management in agriculture can have positive social and economic impacts, which go beyond the achievement of high crop yields, as reported in the sub-theme *Ecosystem services connected to the farm irrigation system*. Water management in forested systems is aimed at a wide range of purposes, including increasing water use efficiency in the Mediterranean forests; prevention of landslide risks; and water quality protection in planted forestry catchments. Innovations in this field are listed in sub-theme *Forest water management*.

Theme 5 is dedicated to integrated supply chain services and tools, innovative farm and forest management. The supply chain theme is very broad as it involves all processes in agriculture and forestry supply chains from planting to end-of-life use. Therefore, a certain degree of overlap with the other themes is expected. The food and wood supply chain consists of many different steps or processes such as planting, fertiliser use, farm and forest management processes, logistics and processing of food and wood and the disposal or reuse of waste and by-products. Sub-themes have been defined to represent the chronological sequence of processes along the supply chain or have been divided into groups of specific processes such as farm management; forest management; management of agroforestry systems; organic farming (and other low input systems); biomass production; precision farming; social innovations (*e.g.* new institutional arrangements, new forms of collaboration); models and decision support systems; natural risks and risk management; logistics; post-harvest technology; certification; quality control and management; wood processing; non-wood forest production; and improved recycling of by-products.

Theme 6 is on *Recycling and smart use of biomass and food waste, in particular waste generated during primary production*. This theme is further specified into *Waste that can be used in primary production*. Wastes can be produced in forests and on farms, and during the processing of food and wood products. Some waste types are quite general and are also covered by other themes. For example, several manure-processing innovations are covered by Theme 3, and wastewater use innovations are largely covered by Theme 4. Theme 6 also contains specific innovations that combine waste processing with energy production (*e.g.* anaerobic digestion, ethanol production, gasification), innovations that process wastes from society into valuable products for agriculture (*e.g.* fertilizers from bones, removal of pollutants from sludge), and innovations that aim to optimize economic and agronomic performance (*e.g.* wood ash recycling, and associated energy production). Recycling of non-organic wastes such as plastic is beyond the scope of this theme. The number of completed mini-factsheets to date (October 2016) is 201 (50, 2, 77, 22, 16, and 34 for Themes 1 to 6, respectively). Appendix Tables 2 and 3 provide two examples of mini-factsheets. Each mini-factsheet has a title, a reference theme and sub-theme, a description of the innovation and deals with one or more issues (problems or processes addressed by the innovation). An issue can be covered by more than one mini-factsheet, meaning that more than one innovation can provide an answer to a given problem.

ask-Valerie.eu

Figures 3, 4 and 5 show an example application of *ask-Valerie.eu*. In Figure 3, the user types a query about *mechanical control*, and the system uses terms from the ontology to assist the user in formulating the query. Figure 4 shows the *term editor*, in which the context of the search terms in the query is presented to better formulate the query. The term editor shows broader, narrower and related terms from the ontology. Narrower terms represent concepts that are *a type of* the term used in the query (*e.g.* the narrower term *mechanical weeding* is a type of *mechanical control*, which was used in the query). The term used in the query is a type of a broader term like *physical control*. Thus, the term editor allows the user to refine the query by including narrower, broader and related terms, or removing terms from the search. Figure 5 shows example results retrieved by the system in response to such a query. Each result is currently represented by a title and a text fragment that represents the section of the document that is most relevant to the user's query. Changing the terms in the term editor prompts *ask-Valerie.eu* to execute another search with the new set of terms, thus empowering the user to access more precise results through stepwise refining of the search.

Discussion

Document base

The document base contains scientific papers, articles in trade journals, practical factsheets, technical manuals, document/reports, and VALERIE mini-factsheets. These documents differ widely in readability, availability, applicability, and language. Research information is often too far removed from practice to be of interest to farmers/advisors and may be more relevant to organisations and companies that aim to develop innovations. We do not know if this latter category of users will benefit from our approach, as some of them may be in direct contact with researchers, especially as more research funding is being available for direct application of research results and multi-actor approaches. Advisors, farmers, and foresters, on the other hand, want practical information on innovations that have been tested and used elsewhere or to guide new innovations. We aim to aggregate a wide array of repositories of such practical information, thereby covering multiple languages. For scientific documents, we aim to give priority to reviews and papers that give examples of interest to farmers and advisors.

To date, we have worked in English to compile the ontology, collect documents, and write the VALERIE mini-factsheets. We are now considering document repositories in other languages than English, as planned in the third and fourth year of the VALERIE project.

Advantages and differences of *ask-Valerie.eu* compared to other search engines

One may ask why we are spending efforts in building a search platform specifically dedicated to agriculture and forestry, when
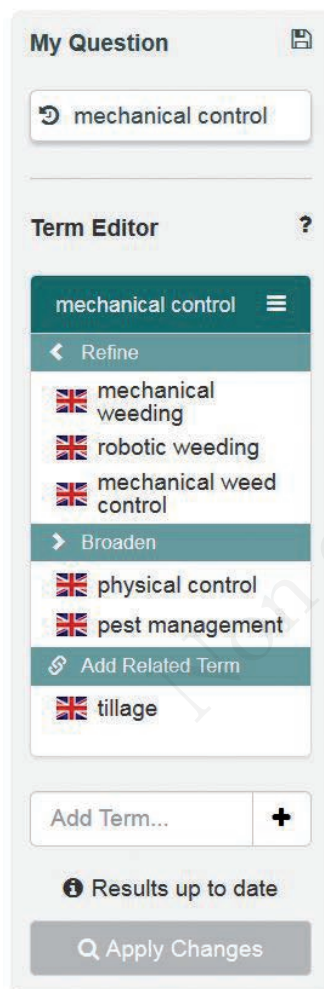

Figure 4. The term editor in which the context of the search terms in the query is shown to better formulate the query: the terms to narrow the query (*mechanical weeding, robotic weeding, mechanical weed control* are three ways to perform mechanical control), to broaden the query (*physical control* and *pest management* are two groups of terms in which *mechanical control* is a member), or to find related terms (mechanical control has to do with *tillage*; the user may also be interested in innovations in this domain).
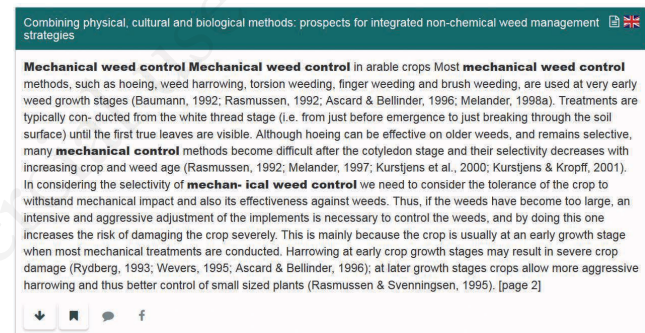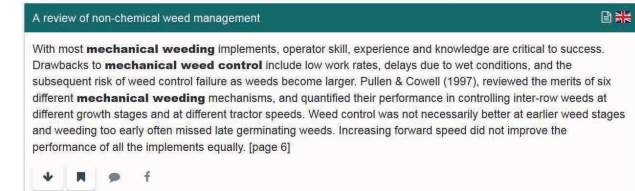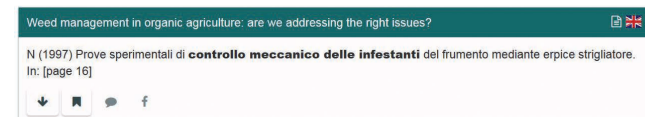

Figure 5. When searching for *mechanical control, ask-Valerie.eu* knows the context of mechanical control. It knows that types of mechanical control, such as *mechanical weed control*, are probably of interest to the reader as well, as results with synonyms of these terms (*e.g. mechanical weeding*), and translations of these terms in the languages that the user understands [*e.g. controllo meccanico delle infestanti* (*mechanical weed control)* if the user understands Italian]. The documents shown are Hatcher and Melander (2003) (A), Bond *et al.* (2003) (B), and Bàrberi (2002) (C).

other search engines are already available and provide documents on these subjects. To answer this question it is useful to list the differences and advantages of our system compared to other search engines (Table 2).

ask-Valerie.eu aims to help farmers and foresters to access relevant knowledge and to get advice about specific innovations that work in their context. This is possible thanks to a consistent input of expert knowledge in *ask-Valerie.eu*, which represents the main difference with other search engines. In selecting relevant repositories, building a dedicated ontology, providing dedicated fact sheets, and evaluating the selection and ranking algorithm, experts insert knowledge that helps users to find best matching answers. Our first steps to do so are to write documents (mini-factsheets), select relevant other documents, and model (annotate) these documents using the VALERIE ontology. Note that we do not use the entire web as the source for document selection; we rather ask experts to indicate which repositories would work best for *ask-Valerie.eu*. Thus, rather than viewing existing engines as a competitor, we consider them as powerful engines to pre-select documents and web pages as sources of information for *ask-Valerie.eu*.

Other differences with current search engines are that *ask-Valerie.eu* enables a smart query-and-answering process (query articulation and expansion; differential ranking of search results based on user profiles) and performs a concept search in which domain knowledge is taken into account. Another difference is that users of *ask-Valerie.eu* can be characterised (geographical location, expertise, *etc*.), and will be able to interact with the system and with other users by writing comments and exchanging information via the search platform.

Finally, by testing the platform with stakeholders in the ten VALERIE case studies, we are learning more about how our users search, and thus will be able to further improve *ask-Valerie.eu* for them. For the development of *ask-Valerie.eu*, we have adopted an iterative approach. In each cycle, new features of *ask-Valerie.eu* are implemented and tested by the case study partners. From these sessions, existing functionality is evaluated, requests for changes and for new features are gathered and prioritised. Then, a new cycle starts. Completion of four cycles in the last two years has led to a clear improvement in the interface, search functionality, language integration, quality of search results and user interaction; thereby enhancing the accessibility of *ask-Valerie.eu* for the target audience. In addition, we envisage that the content and facilities of the platform could be improved in the future through an iterative process of interaction between the users and hosts of *ask-Valerie.eu*

Although other agriculture and forestry-oriented search engines exist, we are not aware of search engines in the agriculture and forestry domains similar to VALERIE. *orgPrints* (http://www.orgprints.org, consulted on 21st October 2016) and *ENDURE* (http://www.endureinformationcentre.eu, consulted on 21st October 2016) are two examples of agriculture-oriented platforms that are dedicated to more specific objectives than VALERIE: *orgPrints* is an archive of documents related to organic farming; *ENDURE* is a document repository dedicated to Integrated Pest Management. The main objectives of *orgPrints* are to communicate research papers and proposals, disseminate research findings, and document research efforts. Users can upload their own documents, browse and search the archive, as well as save and repeat searches. ENDURE offers a search engine that allows to select a crop, pest, control measure, or geographic region from pre-defined lists; then, it shows lists with relevant summaries of articles, reports, videos, sheets, information on books, and links to websites. It also presents – similar to orgPrints – short summaries written by experts. There are many other websites offering concise information to farmers. However, all these examples, although very useful, largely ignore the potential that recent advances in semantic technology may offer. This limits their potential to serve users interactively, and in a manner that is optimally tailored to user needs. In fact, *ask-Valerie.eu* actively uses the previously captured knowledge from the expert to help users of our platform to better formulate their queries and find a better fitting answer.

Ongoing and future work

Future work will focus on both the *ask-Valerie.eu* system and on its building blocks: the ontology and the document base. Work on the ontology involves a number of processes. Firstly, the ontology can be enriched by linking certain subdomains of the ontology, such as *crop*, *weed*, *pest*, *herbicide*, to existing sources. Existing lists of types of crops (http://www.fao.org/fileadmin/templates/ess/documents/world_census_of_agriculture/appendix4_r7.pdf), weeds (http://wssa.net/weed/composite-list-of-weeds/), *etc*. are being added to the VALERIE-ontology. Secondly, the evaluation of *ask-Valerie.eu* by stakeholders will result in the identification of additional practical or colloquial terms that should be added to the ontology. To expand the document base, we are investing in identifying existing repositories of documents, and automatically annotating them. One of the repositories that we are including is CORDIS, which contains documents produced by FP5, FP6, and

Table 2. Comparison of common search engines *vs ask-Valerie.eu*.

Search engines	*ask-Valerie.eu*
Search on literal query, good answers depend on the words you happen to choose	During typing, the system suggests known domain terms to assist the user in finding the best words (Figure 3)
Search on literal query, good answers depend on your skill in using search machines	Query is analysed, domain terms are understood, user is assisted in refining the query (via term editor) (Figure 4)
Literal search, no domain knowledge applied	Concept search in which domain knowledge is taken into account (term + synonyms + narrower terms) (Figure 5)
Aim: find links to webpages that match your search terms	Aim: find answers to the question contained in the shown fragments, document is offered as reference for in-depth knowledge
Search the complete web	Search only preselected document sources matching the *ask-Valerie.eu* domain; the selection has been performed by our domain experts and will in the future be continued by them and by our users
Open system, anonymous users	Community of users

FP7 European projects. The European Union Open Data Portal (https://open-data.europa.eu/en/data/) indicates that 48,860 projects were funded under FP5, FP6 and FP7 calls, in all disciplines. There are about 1500 projects related to agriculture, food, and forestry. We have accessed the CORDIS database to provide access to documents produced by those projects, possibly using also the OpenAire (http://www.openaire.eu) public repository. Other repositories exist outside the EU; in particular, most universities from the US have an extension service delivering practical fact sheets dealing with crop management techniques (*e.g.* Iowa State University, https://store.extension.iastate.edu/; University of Missouri, http://extension.missouri.edu, consulted on 21st October 2016). Substantial effort will be devoted to find documents in non-English languages. We will look for repositories of project deliverables from nationally-funded projects, and for European repositories of factsheets. Examples of such repositories are http://www.infloweb.fr (*To know and manage weeds*, in French; consulted on 21st October 2016), http://www.crpa.it/nqcontent.cfm?a_id=1117 (*CRPA Publications*, in Italian; consulted on 21st October 2016), http://www.forestry.gov.uk/publications (*Forestry Commission/Publications*, in English; consulted on 21st October 2016), and http://www.groenkennisnet.nl (*Green knowledge*, in Dutch; consulted on 21st October 2016). We have already surveyed more than 100 of these repositories and have started contacting their owners. As part of the activities to complete the document base, we plan to prepare a mini-factsheet for most of the innovations on our list.

New documents added to our document base will be automatically screened to check if they describe one of the innovations in our list. This check will verify whether concepts in the new documents match those present in the mini-factsheet (name of the innovation, description of the innovation, and innovation challenges).
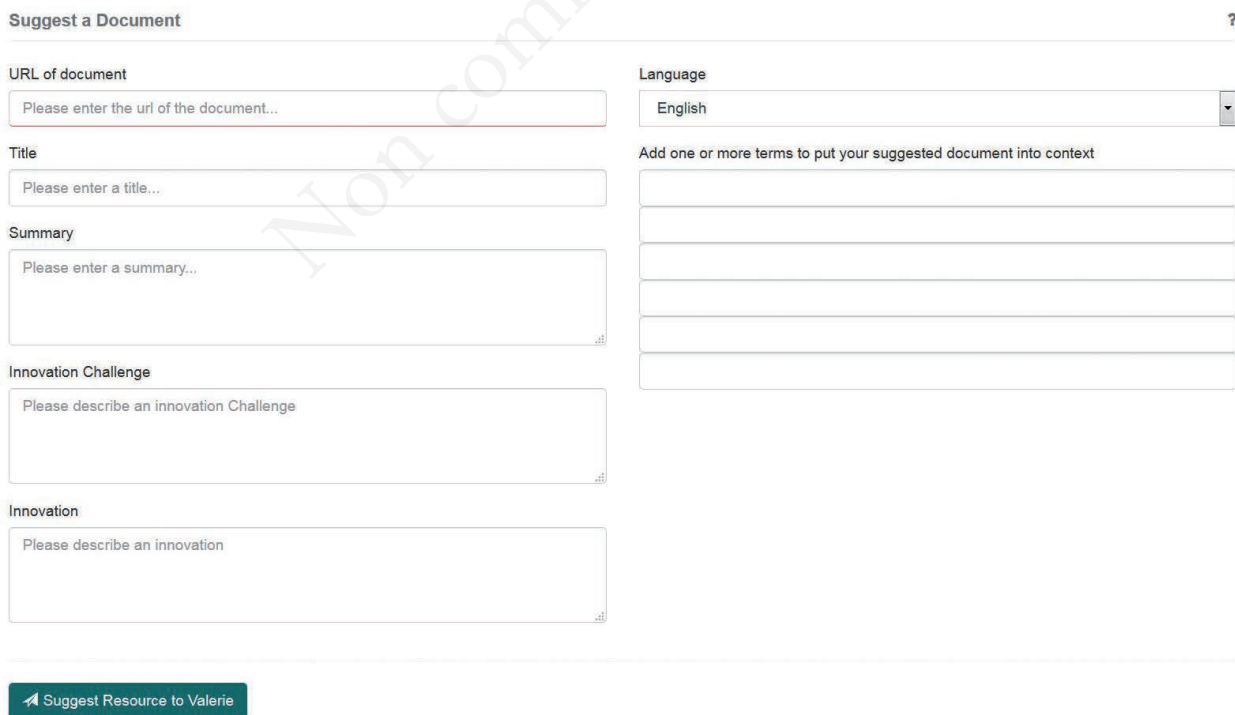
This will also help to overcome the unavoidable subjectivity inherent in the expert-made choice of documents listed in the mini-factsheets.

Future work on *ask-Valerie.eu* will aim to improve the dialogue between the user and the system. To do so, we will study and evaluate different ways of ranking the answers identified by the system. *ask-Valerie.eu* will also be improved by enhancing the user involvement in the system, by allowing users to add their own experiences, their own content such as factsheets or video clips and longer films describing innovative actions, by commenting on the answers provided by the system (indicating what was found useful) and by engaging with each other.

Continuation of *ask-Valerie.eu*

As part of the VALERIE project, actions are being undertaken to maintain *ask-Valerie.eu* and secure its continued availability beyond the lifetime of the project. In particular, the ontology and the document base are components that critically need to be maintained on an on-going basis.

It is inevitable that inconsistencies, incompleteness, and errors will be discovered in the ontology when *ask-Valerie.eu* is widely used. These inaccuracies will have to be fixed. Also, in the project, work has focused on only six thematic domains, with limited attention given to many other domains. Thus, there will be a need to add to the ontology concepts and relations that describe new domains. Finally, concepts that do not exist today may become important in years to come – these will also have to be added. The *ask-Valerie.eu* tool at present has limited functionality for users to point out errors and omissions. This functionality must be expanded. More importantly, a process must be put in place for users to contribute to the system in some organic form, and with some level of moderation, *e.g.* by suggesting documents (Figure 6), innova-

Figure 6. Screenshot of *suggest document* option in *ask-Valerie.eu*.

tions, mini-fact sheets, ontology terms, *etc*.

Curation is an important process not just for the Valerie ontology but for all ontologies. Recently, three important repositories of concepts and terminology in the area of agriculture (AGROVOC, the CAB Thesaurus and the NAL Thesaurus) were combined into one ontology GACS (http://tester-os-kktest.lib.helsinki.fi/ gacsdemo/gacs/en/; consulted on 21st October 2016), which reduces the overall maintenance effort. Conversely, it is possible to create a series of ontologies, each covering a narrow domain, and let each narrow ontology be curated by a group who has the specific expertise and a specific interest to curate. For example, the EPPO Global database (http://gd.eppo.int; consulted on 21st October 2016) provides comprehensive coverage of organisms in agriculture. The future maintenance of *ask.Valerie.eu* will be able to draw upon these methodologies, but will need an organisation or institute to oversee the process. Within the current VALERIE project, we allow users to provide ontology terms and ask dedicated expert groups to integrate these new terms into the existing ontology. When this is done on a regular basis, the amount of time needed is limited.

When new concepts are added to the ontology, all documents in the document base need to be scanned for the new concepts. This is an automated process and like all automated processes, must be monitored and adjusted when incorrect behaviour is detected. Functionality is needed in *ask-Valerie.eu* that allows users to report errors, and experts to address these.

The document base is the second component of *ask-Valerie.eu* that will need curation. We aim to establish a lasting collaboration with aggregators of scientific meta-data (*e.g.* titles and abstracts) and practical documents, with regular updates of their new content as it becomes available to the document base. Work is under way to add output from EC-funded projects, Operational Groups (OGs), Focus Groups (FGs) and so on, as it becomes available. Work will also have to continue on identifying additional repositories of practical information and obtaining permission from their owners to index the content. As with the ontology, the *ask-Valerie.eu* user interface has functionality for users to identify either entire repositories or single documents (Figure 6).

Conclusions

A number of maintenance actions are required – on an ongoing basis – to sustain *ask-Valerie.eu* into the future. This calls for the permanent involvement of organizations that have outreach and knowledge sharing with farmers and advisers as their core business. Most notable among these is the European Innovation Partnership Agricultural Productivity and sustainability (EIP-AGRI) Service Point (http://ec.europa.eu/eip/agriculture/, consulted on 21st October 2016) that aims to embed the *ask-Valerie.eu* tool in their platform hosted by the EC. The sustainability of the platform depends on the provision of annotated documents, updating of the ontology, and hosting of the tool. New materials can be contributed to *ask-Valerie.eu* by European projects dedicated to specific innovation domains, *e.g.* the SmartAKIS thematic network on ICT and robotics for innovation in agriculture (http://www.smart-akis.com/, consulted on 21st October 2016). It could be a requirement of future projects to contribute factsheets; various such options are being explored, including the EIP-AGRI Service Point.

References

Bàrberi P, 2002. Weed management in organic agriculture: are we addressing the right issues? Weed Res. 42:177-93.

Bond W, Turner RJ, Grundy AC, 2003. A review of non-chemical weed management. HDRA, the Organic Organisation, Ryton Organic Gardens, Coventry, UK.

Gruber T, 2009. Ontology. In: L. Liu and M. Tamer Özsu (Eds.) Encyclopedia of database systems. Springer-Verlag, Berlin, Germany, pp. 1-11.

Hatcher PE, Melander B, 2003. Combining physical, cultural and biological methods: prospects for integrated non-chemical weed management strategies. Weed Res. 43:303-22.

Knickel K, Brunori G, Rand S, Proost J, 2009. Towards a better conceptual framework for innovation processes in agriculture and rural development: from linear models to systemic approaches. J. Agric. Educ. Ext. 15:131-46.

Leeuwis C, 2004. Communication for rural innovation: rethinking agricultural extension, 3rd ed. Blackwell Science, Oxford, UK.

Rivera WM, Sulaiman RV, 2009. Extension: object of reform, engine for innovation. Outlook Agric. 38:267-73.

SAI Platform, 2010. Water conservation technical briefs - TB6 - Irrigation scheduling. Available from: www.saiplatform.org/ uploads/Library/Technical%20Brief%206.%20Irrigation%20Scheduling.pdf

Smits RE, Kuhlmann S, Shapira P, 2010. The theory and practice of innovation policy: an international research handbook. Elgar, Cheltenham, UK.

Willems DJM, Koenderink NJJP, Top JL, 2015. From science to practice: bringing innovations to agronomy and forestry. J. Agric. Inf. 6:85-95.