The AGROVOC Concept Server: rationale, goals and usage

Margherita Sini, Boris Lauser, Gauri Salokhe, Johannes Keizer and Stephen Katz
Food and Agriculture Organization of the United Nations (FAO), Rome, Italy

Received 19 October 2007
Reviewed 29 October 2007
Accepted 7 November 2007

Abstract

Purpose – The main objective of the AGROVOC Concept Server (CS) is to create a collaborative reference platform and a "one-stop" shop for a pool of commonly used concepts related to agriculture, containing terms, definitions and relationships between terms in multiple languages derived from various sources. This paper aims to address the issues.

Design/methodology/approach – The CS offers a centralised facility where the agricultural information management community can build and share agricultural knowledge in a collaborative environment.

Findings – The advantages of the CS are its extensibility and modularity that provide the possibility to extend the type of information that can be stored in this system based on user/community needs.

Research limitations/implications – Further investigation still needs to be done on the modularisation of the CS (i.e. the creation of separated ontologies that can still be connected, in order to have domain-related ontologies and to allow for better performance of the CS).

Practical implications – The CS serves as starting point for the development of specific domain ontologies where multilinguality and the localised representation of knowledge are essential issues. Furthermore, it will offer additional services in order to expose the knowledge to be consumed by other applications.

Originality/value – The CS Workbench provides the AGROVOC partners with the possibility to directly and collaboratively edit the AGROVOC CS. It thus provides the opportunity for direct and open "many-to-many" communication links between communities, avoiding decentralised communication between partners and duplication of effort. For the international community, it may allow users to manage, re-use or extend agriculture-related knowledge for better interoperability and for improved services.

Keywords Modelling, Information management, Information systems

Paper type Conceptual paper

Introduction

For over 30 years, the Food and Agriculture Organization of the United Nations (FAO) has facilitated a network of documentation centres from agricultural research and technology institutes and academic faculties. The aim of the network is to enhance the exchange of agricultural knowledge, especially between developing countries. The rise of the Internet in the 1990s has revolutionised the way people share and exchange knowledge. There is little doubt that the web provides a platform for global access to information; however, there are a number of important issues that need to be addressed for this potential to be fully exploited. The web was not initially envisioned as a tool for global access to information, and the underlying standards for information management are not entirely adequate. By the very nature of Internet architecture, information on similar subjects is scattered across many different servers around the world, based on different language, local differences and description.

Library Review
Vol. 57 No. 3, 2008
pp. 200-212

0024-2535
DOI 10.1108/00242530810865745

The views expressed in this publication are those of the authors and do not necessarily reflect the views of the Food and Agriculture Organization of the United Nations.

On the other hand, few tools are available to integrate related information from different sources and, as a result, it is often very difficult to find related information on the web. Such problems can only be solved if action is taken to establish appropriate norms, vocabularies, guidelines and standards to facilitate the integration of data from different sources, and to engage in effective data exchange. Through the adoption of international classification schemes, controlled vocabularies, open standards and common data models, these information management problems shall eventually be overcome. FAO, through the development of tools that exploit such standards, strives to provide an effective framework for "one-stop shopping", where people can search for agricultural information resources without having to explore many different individual websites.

In the agricultural sector, there exist already many well-established and authoritative controlled vocabularies, such as FAOs multilingual thesaurus AGROVOC[1], the CAB Thesaurus[2] and the NAL Thesaurus[3] of the National Agricultural Library in the USA. These systems have been carried over from the traditional library world. However, for these semantic tools to be more effective on the Internet and especially in a multilingual context, there is a need to re-assess the traditional "thesaurus" approach and move towards more powerful models and technologies, such as the development of concept-based systems, known also as "ontologies" (Kang and Lee, 2001; Wielinga *et al.*, 2001; Fisseha and Liang, 2003).

In terms of the Semantic Web, ontology can be defined as a semantic system that describes concepts, the definitions of these concepts, and the specification of relationships amongst them. Ontology takes the traditional thesaurus approach one step further by structuring the terms more formally, by providing richer relationships between concepts and by adding other constraints on concepts that can be exploited by intelligent applications to reason on the concepts and to infer knowledge. Ontologies are an integral part of the Semantic Web, described by Tim Berners-Lee as "an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee *et al.*, 2001).

The AGROVOC Concept Server (CS) project has been conceptualised as a response to this approach and will be used to develop knowledge-based applications that better serve user needs, as already done by others (Clark *et al.*, 2000). The AGROVOC CS is significantly different in that it functions as a resource to help structure and standardise agricultural terminology in multiple languages for use by any number of different users and systems around the world. In addition, the idea is to combine the CS, as a whole or extracted subject ontologies with application ontologies or metadata ontologies, in order to support the semantic infrastructure at FAO (Salokhe *et al.*, 2004; Liang, 2006a).

The idea of the CS

Motivations

The AGROVOC Thesaurus has been successfully used over decades in many systems inside and outside FAO. Its main use has been subject indexing of document like objects in digital libraries in the agricultural domain. In light of the aforementioned emerging technologies, FAO has been investigating the possibility of representing this knowledge in ways that are more suitable for the future web and that can help address the problems of information retrieval and integration (Soergel *et al.*, 2004).

The following points can be summarised as the primary motivations for undertaking the new developments:

- At present, the development and maintenance of the AGROVOC Thesaurus (translation, revision of relationships, suggestions of new terms) is centralised. All work from different partner organisations is directed to FAO for inclusion in the online version of AGROVOC. Given the time lag and cumbersome work carried out to update the online version, the idea was to create a new online version that can be maintained and updated directly by the partners. The distributed and collaborative maintenance would considerably reduce workflow overhead and duplication of effort.
- The current structure of AGROVOC, which follows guidelines for multilingual thesauri, does not allow the users to add localised language information. AGROVOC is English centric meaning that all concepts start from an English version which is then translated into other languages. The idea of the new model is to provide users with the means to depict terms in their own local language, such as common names of plants, and create language specific relationships. This will create a much more powerful linguistic representation of the local knowledge.
- One of the ultimate objectives is to provide better services to users, such as the possibility to retrieve information regardless of the language, spelling or term variants used for searching. Consequently, the introduction of the notion of a concept is required, in which one and the same concept is represented by a set of multiple terms in various languages that all identify the same conceptual idea. Information is hence associated with concepts rather than single terms.
- The traditional and current AGROVOC thesaurus model, albeit useful in many contexts, is limited in its scope. The advances in data modelling, especially using the Ontology Web Language (OWL)[4], offer the possibility to create a concept based structure with explicit semantic relationships and other constraints that can be exploited for automatic inference of knowledge. Offering ways for modularisation, these new models allow the creation of sub-domain ontologies, using only some of the concepts. This allows the users to organise knowledge based on their own distinct application needs.

Compared to the traditional AGROVOC Thesaurus, the main characteristics of the CS are the following:

- It is a concept-based, modularised and extensible system.
- It provides the possibility to realise term and language specific relationships, which offers for much more flexibility on the linguistic level.
- It allows for the representation of more semantics in terms of concept and term relationships and other constraints and definitions provided by the OWL modelling language.
- It accommodates distributed maintenance for improved workflow and better domain coverage.

Currently the CS model[5] is fully defined (Liang, 2006b) and available for reuse. An example of a domain specific ontology – implementing this model for the domain – is the crop wild relatives[6]. The full conversion of the AGROVOC Thesaurus based on this model is under development.

Goals and objectives

With respect to the Semantic Web initiative, the CS strives to:

- provide a framework for sharing common terminology, concept definitions and relations within the agricultural community;
- provide a powerful and extensible model that can be used to create other ontologies;
- streamline efforts and enhance collaboration for the creation of knowledge management systems throughout the agricultural community;
- increase the efficiency and consistency with which multilingual agricultural information objects and resources are described and associated together;
- increase the functionality and relevance when accessing these resources.

Once fully completed and made operational, the CS will offer a contextually rich and modern framework for modelling, serving and managing agricultural terminology. When integrated with web-based search tools, it will greatly facilitate resource retrieval. It should provide access to document-like objects in a variety of languages and offer suggestions for other related resources that are potentially relevant to the topic of interest. The CS is foreseen to empower a variety of useful services such as automatic or semi-automatic translation services, information discovery and reasoning services, guided search services and concept disambiguation services. Such additional functionality will not only dramatically increase the scope of web-based search engines, but also revolutionise the way users interested in agricultural resources interact with the web.

The overall goal is to improve worldwide access to agricultural information and, as such, the CS, in combination with FAO's other standardisation activities, plays a strategic role in its effort to fight hunger with information.

Scope

The CS operates in the agricultural domain (covering agriculture, fisheries, forestry, nutrition etc.) and provides the possibilities to mix-and-match the concepts and their relationships into specific sub-domain ontologies. The maintenance tool of the CS, will assure regular production of AGROVOC as a traditional thesaurus for systems that wish to continue using AGROVOC as a thesaurus. The CS will not necessarily integrate or manage ontologies which are created outside the scope of the CS. However, it may provide links to download these ontologies. The CS will provide web services that can be consumed by application developers to be used to enhance their systems.

Comparison with other approaches

Other models and approaches have already stated the necessity to identify the concept rather than the term as done in traditional thesauri. The ISO Standard 5964[7] (Multilingual Thesauri) clearly states the issues related to term translation. Recently, the British Standard (BS) 8723 has been circulated amongst small expert groups in the areas of SKOS and XML. A thesaurus data model for the BS is already available in UML[8]. Simple knowledge organisation systems (SKOSs)[9] and other similar approaches also take into consideration the notion of the "enhanced thesaurus", so the introduction of concepts vs terms (and more refined relationships) is nothing new to the approach presented here.

Our model has evolved in a long process of research and collaboration using feedback from the agricultural community, so it represents a direct response to the requirements gathered over years from this community and domain. The CS and its maintenance model have been developed with the goal of providing a model that is powerful enough to be extended beyond what thesauri and thesaurus applications can offer. For instance, these approaches introduce the notion of concepts, multilinguality and all the mentioned linguistic advances, but also further constraints (e.g. the possibility for instantiation of concepts and other features that are permitted in OWL). So the strength is in the extensibility. It will always be possible to simplify the model and interface with others such as SKOS, the new British Standard or others.

Moving towards the CS

The CS is an element of the AOS initiative and the core server of knowledge for the new services that FAO, together with partners, would like to offer to its users. The process to move towards the CS has been a gradual one. Over the past few years, FAO and its partners have been carrying out activities that support the transformation of the AGROVOC Thesaurus into the CS.

In particular, the following activities can be pointed out as being crucial.

The idea of the CS first arose in 2001 with the main goal of integrating all the FAO terminological resources. This idea, however, was not very successful because of their disparate structures, scopes and owners. However, the need to restructure AGROVOC as ontology was carried forward. The following figure (Figure 1) depicts the initial vision of the Agricultural Ontology Server (AOS). Back then the AOS was envisioned to be what today is represented by the CS.

In 2002, the first prototype domain ontologies were developed (fisheries, food safety), experimenting with tools to automatically extract concepts and domain

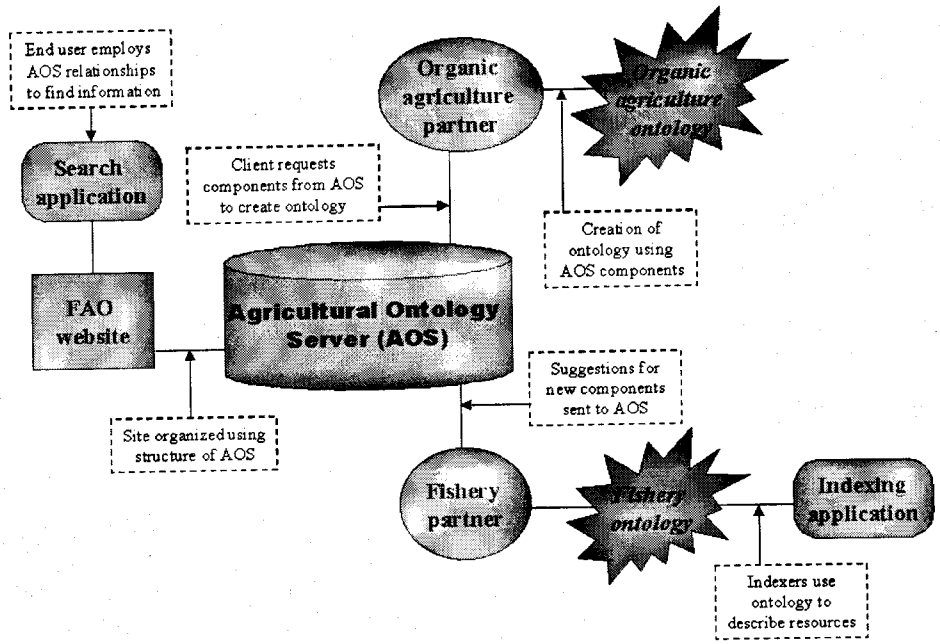

Figure 1.
AOS vision (2001)

specific knowledge of the AGROVOC Thesaurus and using domain experts to validate and develop it further.

Between 2002 and early 2005, a study was conducted to identify the form and technical infrastructure of the CS. The main evaluation was on the possibility of storing an ontology in a relational database (Soergel, 2004).

In 2005, research was carried out to use the OWL for representing the CS structure and data. During this period, OWL was gaining widespread interest from a range of disciplines and domains, including medicine, defence, agriculture, biology and library sciences. More sophisticated and better performing technologies were continually being developed on top of the OWL ontologies. The research provided the following reasons to move from a relational database model to OWL:

(1) OWL allows for easy integration of other RDF-based data sources at the storage level.

(2) OWL allows for straightforward data processing and visualisation.

(3) The OWL model is reusable and interoperable with any RDF triple-store.

(4) OWL is web-enabled which would make data transfer and reuse much more likely. It is also a W3C recommendation.

(5) OWL tools may be reused. For example, reuse and modification of open source tools that use OWL (e.g. Protégé, SWOOP, etc.). These tools already implement many of the functionalities needed for the CS maintenance system.

(6) OWL is an evolving and recognised standard which will assure for maximum interoperability and future support from a wider community.

Subsequently, several efforts were made to represent AGROVOC in SKOS and OWL. In 2006, the AGROVOC CS OWL model was proposed. FAO's research has identified three levels of information, which need to be represented in the CS (Liang *et al.*, 2006b):

- Concept level: a specific notion, concrete or abstract. Every concept will be identified by one or more definitions.
- Term level: any specific sense associated to a concept (every language or synonym is represented as a separate term).
- String level: all word variances representing the same term (here we may include singular and plural forms, spelling variances, abbreviations, etc.).

The power of the semantics comes into place with the possibility to explicitly link concepts with relationships, such as "has pest" or "ingredient of". Such relationships enable the user to learn more about the particular concept and explore the domain around it by following the relationships. Terms may be connected to each other by the identification of exact translations, synonyms or other linguistic information, which may be used to build semi-automatic translation systems. Strings may be differentiated and connected by specifying for example uses of a form (e.g. long forms vs a short one). This new structure allows us to refine and to enhance the information represented originally within a traditional thesaurus.

A revision and refinement of the AGROVOC Thesaurus terms and relationships was carried out, during which many terms were added or revised in order to provide users a better pool of data. All generic thesaurus relationships correspond to multiple meanings (e.g. BT/NT does not necessarily correspond to "subclass of", but to "part of" or any other specific relationship). Therefore, a specific tool was developed to help

experts in revising and refining the thesaurus relationships into more meaningful and more specific semantic relationships. To accommodate linguistic issues like the identification of acronyms, synonyms and other term variances, or specific terms such as all chemical-related terms, a revision of the term relationships (mainly between descriptors and non-descriptors) and scope[10] was necessary. This activity is only partially completed and currently ongoing. Furthermore, better organisation of scientific and common names for taxonomic entities is necessary. The current AGROVOC Thesaurus does not clearly distinguish between taxonomic terms and common names; they may be related to each other with the "Related Term" relationship, but may not indicate that they really describe the same concept.

Currently, the development of a tool for the collaborative maintenance of the data pool is ongoing (CS Workbench). A preliminary version is available for demonstration and is undergoing testing[11].

In parallel, several projects have been undertaken to extract specific sub-domain ontologies from AGROVOC for use in specific information systems or applications, both inside and outside FAO. These include:

- Food safety ontology (Lauser *et al.*, 2002; Volz *et al.*, 2003)
- Food, Nutrition and Agriculture portal (Sini *et al.*, 2007b)
- Fishery ontology (Gangemi *et al.*, 2002; Gangemi *et al.*, 2004)
- AGROVOC Topic Map (Kawtrakul *et al.*, 2007)
- Fertilizers ontology (Sanchez-Alonso and Sicilia, 2007)
- Crop Wild Relative ontology (Morten, 2007)

The CS and its role in AOS today

Over the years, the initial idea of the AOS developed into something much larger. The Agricultural Ontology Service, of which the CS is now an integral part, also includes domain ontologies, registries of mappings, URN services to name but few. Figure 2 below depicts the overall architecture of the "new" Agricultural Ontology Service. It hosts a wide variety of elements and services which are necessary to realise interoperability in the agricultural domain and which will be made available to users in the international community for better harmonisation of data and better tools development. In Figure 2:

(1) The CS as described above represents the core of the AOS.

(2) A knowledge organisation system (KOS) registry will be maintained in order to register trusted and well-developed KOS within the agricultural community (whether based on the CS model or created using other models).

(3) A registry of mappings will be maintained through which mappings between featured KOS will be made available for use to be incorporated into other systems for disambiguation, translation and other purposes.

(4) Other ontologies and KOS developed under the umbrella of the AOS (Fisheries Ontologies, Nutrition, Crop Wild Relatives, etc.) will be made available through AOS.

Additional, AOS Services will include (but are not limited to):

- a URN registry;

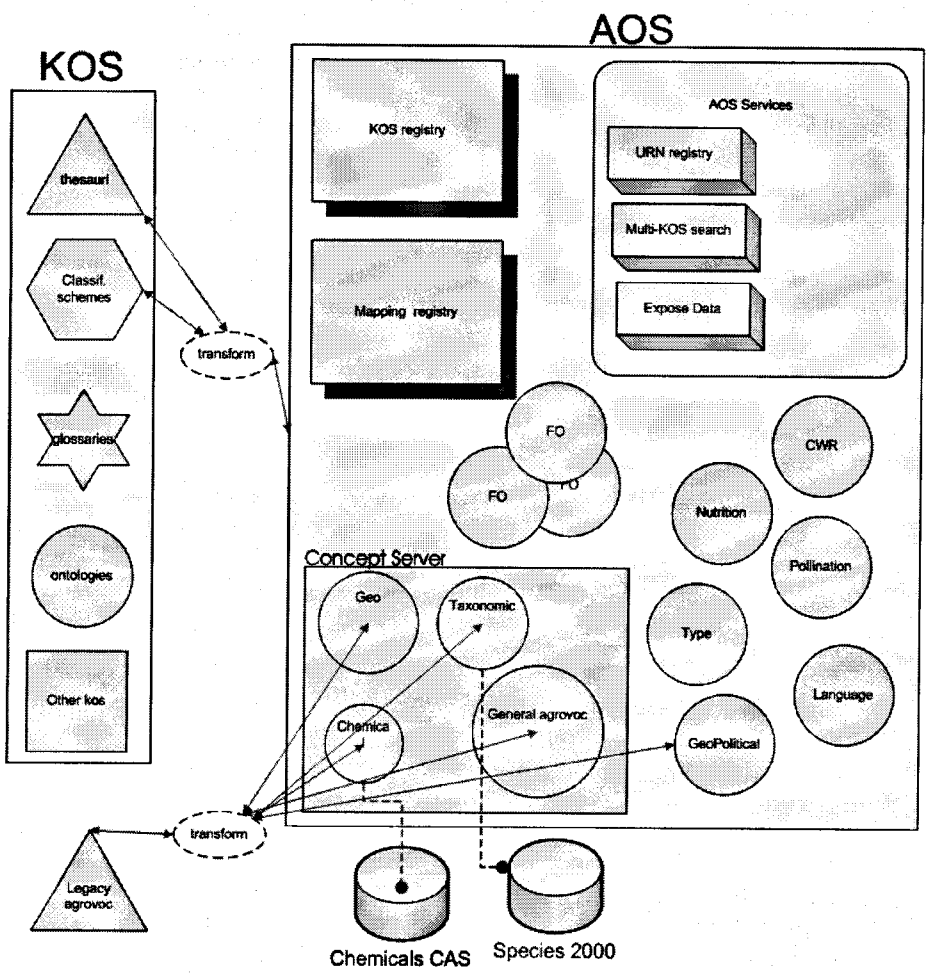

Figure 2.
Agricultural Ontology Service vision (2007) and the role of CS within

- a relationship registry for approved relationships within the AOS and agricultural community;
- multi-KOS search for keyword and semantic searches across several KOS, exploiting mappings and other features of the model;
- web services to access hierarchies, concept details, linguistic information, relationships, etc. of the CS or other featured KOS.

Community and roles

The AGROVOC Thesaurus is currently used and maintained by different partner organisations around the world. However, there is lack of synchronisation in content creation and updating, and there are only generic and poor common guidelines for standard actions or good practice (Sini *et al.*, 2007a). Incorporating local knowledge and other languages is therefore cumbersome and slow.

Figure 3 shows the current workflow for updating AGROVOC and managing translations and updates in a centralised environment.

The CS Workbench is a collaborative tool for ontology management. It can be used by different users who may have different privileges and roles. The different roles that have been identified and are currently implemented are as follows:

(1) Not logged-in users: These users can browse or search the system and submit suggestions for terms in any language, concepts and relationships. They can also provide other comments and feedback on the system in general.

(2) Term editors (terminologist/thesaurus editor): Term editors are content experts in specific domains related to agriculture. They have full permissions on the term level in their assigned languages but have no rights for concepts modifications; however, they can still make suggestions for concept modifications.

(3) Ontology editors (more experienced terminologist/thesaurus editor): Ontology editors are experienced in ontology modelling and are familiar with the concept-term-string level. They have full permissions to manage concepts and terms in their assigned languages.

(4) Validators: Validators are experts that can check and validate the work done by the editors (terms and ontology editors) depending on their assigned languages. They should be experienced both in agricultural content and ontology modelling practices.

(5) Publishers (ontology editors): Publishers have full permission on terms and concepts in all languages. They generally confirm the work of the validators, but can directly validate ontology editors and term editors also. They are the final instance in the validation workflow.

(6) Administrators: Administrators have full access to all system functionalities and all languages. They will manage users, groups and permissions, provide statistics and perform other duties.

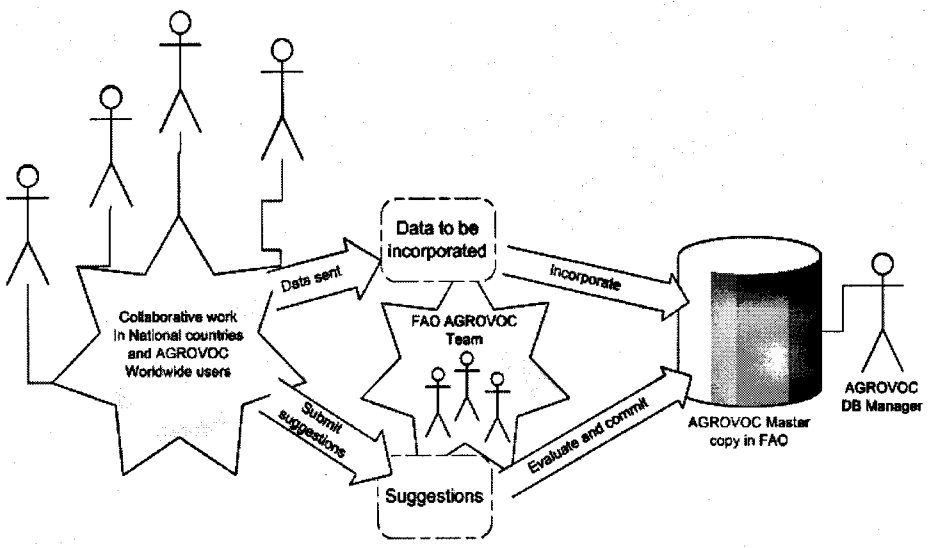

Figure 3.
Current workflows for translations and maintenance of AGROVOC

The current AGROVOC partners will be contacted to serve the different roles as individuals or as an organisation. This new collaborative infrastructure will significantly change the current workflow drawn in Figure 3. The new infrastructure proposes a system, in which all actors will interact collaboratively and concurrently (Figure 4).

The collaboration in this case is much more effective because:

- AGROVOC editors all over the world can have direct access to the maintenance tool;
- changes are immediate and there is no need to wait for FAO actions (apart from the validation phase which will be carried out by FAO);
- all users can immediately see and benefit from users' contributions;
- the cycle of adding data to AGROVOC and reusing it in the respective systems is more immediate; after data is inserted in the system, and eventually validated, it is immediately available for remote access through web services or for download in various formats.

Owing to the collaborative nature of the CS and the number of people that will eventually interact via the new infrastructure, well defined and managed workflows are necessary to avoid confusion, data inconsistency and to ensure quality control. A validation workflow and user rights management system is currently being implemented in the CS Workbench. For example, only Administrators will have the right to edit, add or delete the list of relationships that can be used to link concepts and terms.

Intellectual property rights (IPR) and data custody issues

The IPR, in particular copyright, of material such as terminological data, glossaries, images, and so forth, shall remain with the originating party, who will be indicated as the source partner if the information is reproduced or disseminated through the CS or elsewhere. Copyright of the information, as well as rights to any other intellectual

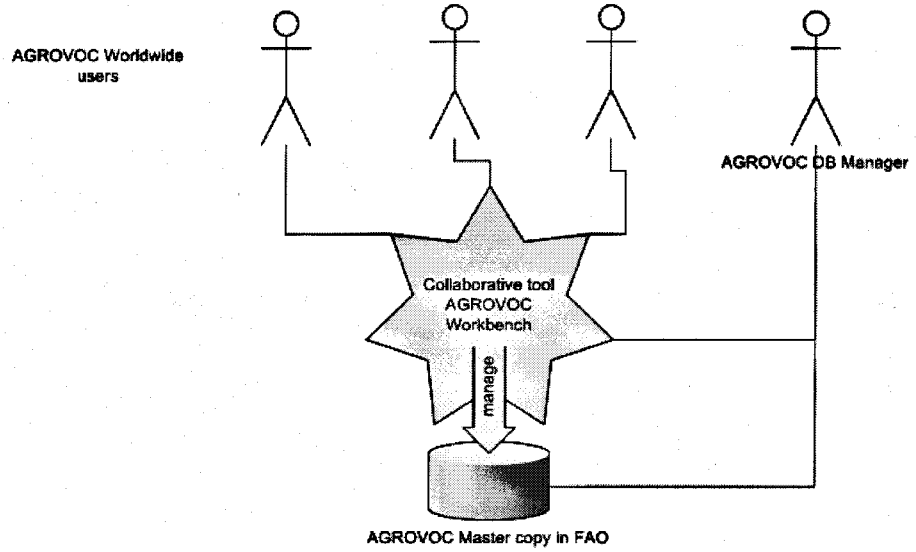

Figure 4.
Proposed infrastructure for collaborative CS management and use

property, developed in collaboration within the network shall be jointly vested to all parties involved. Each party shall have full rights to exploit such jointly owned works after informing the other parties, without the need of approval from the other. All partners to the network shall have free access to all information developed. As the IPR issues are extremely important to ensure that work is rightfully recognised, this area will be further discussed by the partners in the next few months of development.

Conclusion and future work

The CS is conceived as a pool of semantically related concepts. All concepts are represented with multiple terms and definitions in many languages. Multilinguality is addressed in a very complete but extensible conceptual model. The model is capable of arranging the complex multilinguality as managed by FAO (even accounting for non-Latin languages) and terminology information. Terms have relationships that allow the identification of exact translations of terms, synonym relationships or word variances. A major advantage of the CS is its extensibility: The amount and type of information that can be stored in this system can be extended anytime based on needs. For example, at the term level, we could add information about the term etymology, which will allow AOS to provide additional services to specific users. Other users could take the CS or parts of it as a basis and extend the model to make it "OWL Full". In this way, users would be able to treat concepts as instances and express even more sophisticated knowledge according to their needs.

The CS is one of the key elements of the AOS initiative, which aims to provide better services to users of the agricultural domain exploiting new semantic technologies. The full AOS initiative will provide additional tools and services for the exploitation of the data contained in the CS (e.g. tools for semi-automatic generation of taxonomies or domain ontologies to be used as a basis for new FAO websites, web services for accessing CS data, tools for semi-automatic translation, for semi-automatic indexing, etc.).

The full AGROVOC Thesaurus has been converted to a unique ontology, which can be further modularised by top concepts, categories or classification schemes. However, further investigation is required with respect to the modularisation of the CS (i.e. the creation of separated ontologies that can still be connected, in order to have domain-related ontologies and to allow for better performance of the CS).

Notes

1. AGROVOC Thesaurus: www.fao.org/aims/ (accessed 18 December 2007).
2. CAB Thesaurus: www.cabi.org/DatabaseSearchTools.asp?PID = 277 (accessed 18 December 2007).
3. NAL Thesaurus: http://agclass.nal.usda.gov/dne/search.shtml (accessed 18 December 2007).
4. Ontology Web Language www.w3.org/2007/OWL/wiki/OWL_Working_Group (accessed 18 December 2007).
5. Concept Server Model: www.fao.org/aims/aos/aos.owl (accessed 18 December 2007).
6. Crop Wild Relatives Ontology: www.fao.org/aims/aos/cwr_DL.owl (accessed 18 December 2007).
7. ISO 5964:1985 Documentation – Guidelines for the establishment and development of multilingual thesauri www.collectionscanada.ca/iso/tc46sc9/standard/5964e.htm (accessed 18 December 2007).
8. http://isegserv.itd.rl.ac.uk/blogs/alistair/archives/40 (accessed 18 December 2007).

9. Simple knowledge organisation systems (SKOS): www.w3.org/2004/02/skos/ (accessed 18 December 2007).
10. Scope is specific information of the AGROVOC thesaurus identifying if a term belong to a specific sub-vocabulary such as geographical terms, taxonomical terms or chemicals.
11. http://vivaldi.cpe.ku.ac.th:8085/agrovoc/ (accessed 18 December 2007).

References

Berners-Lee, T., Hendler, J.A. and Lassila, O. (2001), "The Semantic Web", *Scientific American*, May 2001, available at: www.sciam.com/article.cfm?articleID = 00048144-10D2-1C70-84A9809EC588EF21 (accessed 18 December 2007).

Clark, P., Thompson, J., Holmback, H. and Duncan, L. (2000), "Exploiting a thesaurus based semantic net for knowledge-based search", in *Proceedings of IAAI-2000*, available at: http://citeseer.ist.psu.edu/309406.html (accessed 18 December 2007).

Fisseha, F. and Liang, A.C. (2003), "Reengineering AGROVOC to ontologies: steps towards better semantic structure, *NKOS Workshop*, 31 May 2003, Rice University, Houston, available at: ftp://ftp.fao.org/gi/gil/gilws/aims/publications/presentations/2003_7.pdf (accessed 18 December 2007).

Gangemi, A., Fisseha, F., Keizer J., Liang, A., Pettman, I., Sini, M. and Taconet, M. (2004), "A core ontology of fishery and its use in the fishery ontology service project", in *Proceedings of the EKAW*04 Workshop on Core Ontologies in Ontology Engineering, 8 October, Northamptonshire, UK*, available at: http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-118/paper4.pdf (accessed 18 December 2007).

Gangemi, A., Fisseha, F., Pettman, I. and Keizer, J. (2002), "Building an integrated formal ontology for semantic interoperability in the fishery domain", *Agricultural Information and Knowledge Management Papers*, available at: ftp://ftp.fao.org/docrep/fao/008/af242e/af242e00.pdf (accessed 18 December 2007).

Hulden, M. (2007), "A practical approach on creating a restricted ontology for crop wild relatives", *Agricultural Information and Knowledge Management Papers*, available at: ftp://ftp.fao.org/docrep/fao/010/ah862e/ah862e.pdf (accessed 18 December 2007).

Kang, S.J. and Lee, J.H. (2001), "Semi-automatic practical ontology construction by using a thesaurus, computational dictionaries, and large corpora", in *Proceedings of the Workshop on Human Language Technology and Knowledge Management (Annual meeting of the ACL)*, available at: http://portal.acm.org/citation.cfm?id = 1118226 (accessed 18 December 2007).

Kawtrakul, A., Yingsaeree, C. and Andres, F. (2007), "A framework of nlp based information tracking and related knowledge organizing with topic maps", in *Natural Language Processing and Information Systems* (*Lecture Notes in Computer Science*), SpringerLink, Berlin/Heidelberg, pp. 272-83.

Lauser, B., Wildemann, T., Poulos, A., Fisseha, F., Keizer, J. and Katz, S. (2002), "A comprehensive framework for building multilingual domain ontologies: creating a prototype biosecurity", in *Proceedings of the International Conference DC-2002: Metadata for e-Communities: "Supporting Diversity and Convergence", 13-17 October*, Biblioteca Nazionale Centrale Firenze, available at: www.bncf.net/dc2002/program/ft/paper13.pdf (accessed 18 December 2007).

Liang, A., Salokhe, G., Sini, M. and Keizer, J. (2006a), "Towards an infrastructure for semantic applications: methodologies for semantic integration of heterogeneous resources", *Cataloging and Classification Quarterly*, Vol. 43 Nos. 3-4, pp. 161-89, available at: ftp://ftp.fao.org/docrep/fao/009/ag869e/ag869e00.pdf (accessed 18 December 2007).

Liang, A., Lauser, B., Sini, M., Keizer, J. and Katz, S. (2006b), "From AGROVOC to the agricultural ontology service/concept server: An OWL model for managing ontologies in the

agricultural domain", in *Proceedings of OWL: Experiences and Directions Workshop, 10-11 November*, Atlanta, available at: http://owl-workshop.man.ac.uk/acceptedPosition/submission_31.pdf (accessed 18 December 2007).

Salokhe, G., Pastore, A., Richards, B., Weatherley, S., Aubert, A., Keizer, J., Nadeau, A., Katz, S., Rudgard, S. and Mangstl, A. (2004), "FAO's role in information management and dissemination – challenges, innovation, success, lessons learned", *Quarterly Bulletin of the International Association of Agricultural Information Specialists (IAALD)*, Vol. 49 Nos. 3-4, pp. 73-83, available at: www.fao.org/docrep/008/af238e/af238e00.htm (accessed 18 December 2007).

Sanchez-Alonso, S. and Sicilia, M.A. (2007), "Using an AGROVOC-based ontology for the description of learning resources on organic agriculture", in *Proceedings of the Second International Conference on Metadata and Semantics Research (MTSR'07), Corfu, Greece, 11-12 October*, available at: www.cc.uah.es/ssalonso/papers/AGOVOC_mtsr.pdf (accessed 18 December 2007).

Sini, M., Johannsen, G. and Salokhe G. (2007a), *Basic Guidelines for Managing AGROVOC*, FAO, Rome, available at: ftp://ftp.fao.org/docrep/fao/010/ai144e/ai144e00.pdf (accessed 18 December 2007).

Sini, M., Salokhe,G., Pardy, C., Albert, J., Keizer, J. and Katz, S. (2007b), "Ontology-based navigation of bibliographic metadata: example of the Food, Nutrition and Agriculture Journal", in *Proceedings of the International Conference on the Semantic Web and Digital Libraries (ICSD-2007), 21-23 February, DRTC, Bangalore*, available at: ftp://ftp.fao.org/docrep/fao/009/ah765e/ah765e00.pdf (accessed 18 December 2007).

Soergel, D., Lauser, B., Liang, A., Fisseha, F., Keizer, J., Katz, S. (2004), "Reengineering thesauri for new applications: the AGROVOC example", *Journal of Digital Information*, Vol. 4 No. 4, available at: http://jodi.tamu.edu/Articles/v04/i04/Soergel/ (accessed 18 December 2007).

Volz, R., Studer, R., Maedche, A. and Lauser, B. (2003), "Pruning-based identification of domain ontologies", in *Proceedings of I-KNOW '03, 2-4 July, Graz, Austria*, available at: http://i-know.know-center.tugraz.at/previous/i-know03/papers/kc/volz.pdf (accessed 18 December 2007).

Wielinga, B.J., Schreiber, A. Th., Wielemaker, J. and Sandberg, J.A.C. (2001), "From thesaurus to ontology", in *Proceedings of the First international conference on Knowledge capture, ACM Press, Victoria, British Columbia*, pp. 194-201, available at: www.cs.vu.nl/~guus/papers/Wielinga01a.pdf (accessed 18 December 2007).

Corresponding author
Margherita Sini can be contacted at: margherita.sini@fao.org