

Basic Guidelines for Managing AGROVOC

Prepared by: Margherita Sini, Gudrun Johannsen, Gauri Salokhe
{Firstname.Lastname@fao.org}

With the contribution of: Anita Liang, Dagobert Soergel
Last revision: April 2008

Basic Guidelines for Managing AGROVOC	1
1. Introduction	2
2. Resources for translations and maintenance	2
3. Guidelines for translations	3
4. Adding new terms to AGROVOC	4
4.1. General conventions	5
5. Managing relationships in AGROVOC	6
5.1. Manage relationships using the maintenance tool.....	7
5.2. Manage relationships using the refinement tool	10
6. AGROVOC structure and content revision.....	13
7. AGROVOC online	14
8. Accessing AGROVOC from your applications/systems	15
9. AGROVOC and copyrights	15

1. Introduction

AGROVOC is a Multilingual Agricultural Thesaurus developed by FAO. It is an agricultural thesaurus in its broadest sense, covering not only agriculture but also fisheries, forestry, human nutrition, environment, and relevant terminology in the supporting biological, physical, and social sciences. AGROVOC was basically "born" in English, or at least was (and still is) oriented to Western-European languages (mainly EN, ES, FR). However, over the years, the need of translating it in multiple languages arose, in particular in the non-European Chinese and Arabic, as these are official FAO languages. And the translations were later expanded to many other languages. This is the current situation (April 2008):

- AGROVOC is available online in seventeen languages (English, French, Spanish, Arabic, Chinese, Japanese, Portuguese, Thai, Czech, Slovak, German, Hungarian, Polish, Farsi/Persian, Italian, Hindi and Lao) from: http://www.fao.org/aims/ag_alpha.htm;
- it is under development for: Russian, Moldavian, Turkish, Telegu, Marathi and other Indian languages; minor efforts are ongoing for Korean, and Malay;
- we know people have expressed interest for translation into: Amharic and Catalan.

Updates to the thesaurus are published online and in the ftp area every 3 months.

This document can be used by AGROVOC managers, FAO partners, and other AGROVOC users. It may be updated over the years.

2. Resources for translations and maintenance

AGROVOC has been first created by domain experts in agriculture. Also the translations are generally executed by domain experts in the native language we want to translate it into. Experts, for their work of translating terms, can either use a specific FAO resource called FAOTERM (<http://www.fao.org/faoterm>) which is a terminology database with translations in five languages, or examine agricultural resources in order to see the current and real use of a term (in general you can refer to AGRIS resources, but the Web is also a valid resource).

Other resources for translation and enrichment are available from:

- http://www.fao.org/aims/kos_list_dom.htm
- <http://www.fao.org/faoterm/map.asp?lang=EN&open9=1&what=2>

There are several guidelines for building thesauri:

- Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies, ANSI/ISO Z39.19-2005;
- Guidelines for the establishment and development of multilingual thesauri, ANSI/ISO 5964-1985;
- Guidelines for Multilingual Thesauri, Working Group on Guidelines for Multilingual Thesauri, Classification and Indexing Section, IFLA, April 2005;
- The FAO House style: <http://www.fao.org/docrep/004/AC339e/AC339E00.htm>.

3. Guidelines for translations

Generally, the reason for translating AGROVOC is the need for its use in specific countries (usually for indexing by AGRIS centres¹). Generally, in these countries, a specific Institution or Ministry² downloads the digital copy of the thesaurus and performs the translations of terms to be used in indexing systems. Of course, in general, the translation of every term is done taking into consideration its definition (confirmed further by examining all the broader terms, the narrow terms, the related terms, scope notes, etc.). The translated versions of AGROVOC in these languages are then sent to FAO, which takes care of incorporating it all together in a master copy.

A valid alternative method would be to install locally the Web-based AGROVOC Maintenance tool. This tool is available for download from sourceforge.net³. If you decide to proceed this way, feel free to download it, install it, and test it. Once you are confident with it, FAO can send you the latest version and can provide you the latest copy of the database in order to translate it.

The maintenance tool uses the traditional thesaurus relationships. However, it will be good to refine the AGROVOC relationships whenever possible (see Chapter 5). For this task you may use the refinement tool in order to ensure more correctness and consistency between terms in AGROVOC with the final objective of improving the quality of the AGROVOC thesaurus. The refinement tool can also be downloaded from sourceforge.net.

When translating AGROVOC keep in mind the following:

1. Consider not only the labels (terms) in various languages, but consider also all relationships and scope notes; this will help you understand the real meaning of a term.
2. Translate first all descriptors. Those are the ones that are used for indexing.
3. Feel free to add new terms in your language if you think it is appropriate. Assign at least one English translation or an English definition (use the scope fields). (See also next chapter)
4. When adding new descriptors to AGROVOC, insert them at the proper place in the hierarchy by assigning the most specific broader term (BT).
5. Feel free to translate also non-descriptors which do not exist in English or other languages. If the translation of a non-descriptor in your language is the same as the translation of the corresponding descriptor, then DO NOT repeat the translation, and keep only the descriptor. In case a descriptor has no exact translation but the corresponding non-descriptor has exact translation, then use the translation of the non-descriptor term as the preferred term.
6. If a term has more than one translation or has synonyms or spelling variants in your specific language, create a new non-descriptor as follows:
 - 6a. First check if non-descriptors already exist in other languages (for example in English, French, Spanish, etc.): if you can identify that a synonym or spelling

¹ <http://www.fao.org/agris/>

² http://www.fao.org/agris/Centre.asp?Content=NW&Menu_1ID=NW&Menu_2ID=NW1&Language=EN

³ AGROVOC Maintenance Tool <http://sourceforge.net/projects/agrovoc/>

variant in your language corresponds to one of them, then just add your translation to the existing non-descriptor, otherwise create a new term (point 6b).

- 6b. If there is no non-descriptors corresponding to your language, create a new term as non-descriptor (the system will assign a new term code), and link it to the corresponding descriptor with the relationships (USE/USED FOR).
7. For new terms, try to create simple concepts and non compound terms. E.g. instead of creating “Agricultural and rural legislation”, create instead two terms “Agricultural legislation” and “Rural legislation”.
8. If possible use the refinement tool to refine AGROVOC relationships.

Report any problem you may have with the tool to FAO (FAO-AGRIS-CARIS@fao.org).

4. Adding new terms to AGROVOC

Whenever a concept is missing in the thesaurus, a new descriptor must be added, following the rules given below to decide which term should be included into AGROVOC and in which way.

1. Check in dictionaries, glossaries, other thesauri, such as CAB Thesaurus or NAL Agricultural Thesaurus (<http://agclass.nal.usda.gov/agt/agt.htm>), catalogues, on the Web, etc., if the term exists, and check for the most used term (spelling, etc.). A word may have more than one spelling, e.g. *labour* - *labor*; *aetiology* - *etiology*, *centre* – *center*, etc.
2. Check if a scope note / definition might be useful in order to ensure the correct use of the new descriptor.
3. For names of animals and plants, look for taxonomic names and common/local names. Examples:
Apis mellifera – *Apis mellifica*, *honey bee*, *honeybee*;
Vicia faba – *broad bean*, *faba bean*, *field bean*, *horse bean*.
(decide which one will be descriptor and which ones will be non-descriptors).
There is no rule for making a scientific name or a common name descriptor or non-descriptor. Both types of terms can be descriptor or non-descriptor based on the frequency one is used. More recently FAO is adopting a policy to make them both descriptors and link them with the RT relationship.
4. For names of countries, use the official names as defined in the FAO Terminology <http://www.fao.org/faoterm/nocs.asp?lang=EN>, and/or in the UNBIS (United Nations Bibliographic Information System) thesaurus [http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/\\$\\$searche?OpenForm](http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/$$searche?OpenForm).
Generally all countries recognized by the UN are already in AGROVOC. As AGROVOC geographical entities follow a specific UN standard, just limit to add the translations for existing geographical entities. You can also add geographical places below country level.

Short names or abbreviations/acronyms may be entered as synonyms, e.g.

Slovak Republic - *Slovakia*

European Union – *EU*

(decide which one will be descriptor and which ones will be non-descriptors).

5. For entering a new term into AGROVOC, please follow the *ISO standard 5964:1985, Documentation-Guidelines for the establishment and development of multilingual thesauri*, as mentioned in Chapter 2. If possible, terms should be nouns or noun phrases, but in some multiword terms, phrases may be used. Avoid commas, periods, and hyphens.

4.1. General conventions

Number

The form of the English terms in AGROVOC follows the Monolingual Controlled Vocabularies ISO standard (e.g. all countable terms should be plural, non-countable terms should be singular, etc.), but the French version of AGROVOC follows a different methodology, and so may other languages.

However, in the context of making AGROVOC a richer base of terminology, we are open to creating nouns in singular form.

Therefore, nouns may also be in the singular except of

- pluralia tantum, i.e. terms that appear only in the plural and do not have a singular;
Examples: *clothes, scissors*;
- proper nouns occurring in the plural;
Example: *United Nations*;
- special terminologies;
Example: *Chordates*.

Case

As AGROVOC was originally built following the mentioned standards, generally all common nouns in the thesaurus now start with an initial capital. However, we also accept lower case, as this is a preferable form.

Examples: *agriculture; rural development*.

Exceptions:

- All non-function words (i.e. substantives, verbs, adjectives) occurring in proper noun phrases should begin with an upper case letter.
Example: *Union of Banana Exporting Countries*.
- Also country names and other geographic names should begin with an upper case letter.
Examples: *Italy, Papua New Guinea, Atlantic Ocean, Lake Victoria*.
- Acronyms should generally appear in upper case.
Examples: *FAO, UNICEF*.

Homographs

Homographs are terms that have the same spelling as another expression but represent two or more different concepts. In these cases the meaning is clarified by the use of a parenthetical qualifier that follows the term.

Examples: *Alabama (Lepidoptera), Alabama (USA)*.

Scope notes

A scope note serves to explain the meaning and application of a term.

Examples:

Application rate (Scope note: *Of drugs, fertilizers, pesticides or other agricultural chemicals*);

Equipment (Scope note: *Use only when a more specific term is not available*);

A scope note can also be used for:

- cross-references to terms representing other concepts than that chosen for the term under consideration:

Example: *Dehydration* (Scope note: *Excessive loss of water due to physiological phenomena; for the process of removing water from materials or products use “drying”*); NOTE: in this case do not enter the cross-referenced term, but enter his term code together with “<>” characters (e.g. “*Excessive loss of water due to physiological phenomena; for the process of removing water from materials or products use <2402>*”).

Note: This is just an internal representation of the terms: the AGROVOC tools take this in consideration by just replacing the term code with the appropriate term.

- indicate limitations or extensions of the meaning of the term:

Examples:

Consistency (Scope note: *Restricted to the physical property*);

Agricultural sector (Scope note: *Includes fishery and forestry sectors*);

Growth rate (Scope note: *Restricted to the biological phenomenon; in economics use “Economic growth rate”*) – (also in this case reference the other term with his term code: “*Restricted to the biological phenomenon; in economics use <29767>*”). Even better the term “*biological growth rate*” could be created.

Definitions:

With the maintenance system you may also insert definitions.

Example: *Exclusive economic zones* (Definition: *Areas beyond the territorial waters in which marine resources fall under the jurisdiction of the coastal states*);

5. Managing relationships in AGROVOC

There are two possibilities for managing relationships in AGROVOC:

1. Using the maintenance tool: this will allow the creation of new relationships or delete existing relationships;
2. Using the refinement tool: this will also allow the modification (refinement) of already existing relationships.

5.1. Manage relationships using the maintenance tool

The maintenance tool appears with a first page asking for a login and password:

Access to Agrovoc Maintenance Interface

Agrovoc
an application to browse the agrovoc database

January 2006

Please Login

Interface Language:

User name:

Password:

Or use the [browsing interface](#) | [Refinement Tool](#)

After logging in you are allowed to change the content of the database using the interface. The main menu proposes several functionalities such as:

Manage Terms

- Search Terms
- Alphabetical List
- Add a new Term
- Edit a Term
- Word Tree
- Browse Hierarchy
- What's new
- Modify scope note and Descriptions
- Delete one or more Term/s
- Browsing Categories

The management of relationships is done via the “Word Tree” function, but using any other function (e.g. search) you may go to the “Word Tree” using this button:

Once you are in the “Word Tree” page, you need to have selected one term, in order to see or modify its relationships. In the following example, the term “cows” is selected:

The screenshot shows the Agrovoc Word Tree interface. At the top left, it says "Agrovoc an application to browse the agrovoc database". At the top right, it says "Word Tree". Below this is a navigation menu with links: Home | Search | Alphabetical List | Hierarchy | News | Add a Term | Edit | Delete Terms | Word Tree | Refinement Tool | Edit scope notes and Descriptions | Categories | Help | Logout (You are login as: GILW).

The main content area contains a table with the following structure:

Relationship types		Date: 2006-06-13 17:24:47
Language: English	Term: Cows (1939) Descriptor with relations	Search: <input type="text"/> Find Term
Type of relationship: Broader Term (50)		help
Language: English	Related to: No term selected.	Search: <input type="text"/> Find Related Term

If no term is selected you may enter a term or part of the term and select “Find Term” to search for it:

Relationship types	Date: 2006-06-13 17:27:23	
Language: <input type="text" value="English"/>	Term: No term selected.	
	Search: <input type="text" value="water"/>	<input type="button" value="Find Term"/>
Type of relationship: <input type="text" value="Broader Term (50)"/>		help
Language: <input type="text" value="English"/>	Related to: No term selected.	
	Search: <input type="text"/>	<input type="button" value="Find Related Term"/>

8309	+ Water	<input type="button" value="SELECT"/>
24929	+ Water activity	<input type="button" value="SELECT"/>
16064	+ Water analysis	<input type="button" value="SELECT"/>
8310	+ Water availability	<input type="button" value="SELECT"/>
8311	+ Water balance	<input type="button" value="SELECT"/>
8312	+ Water binding capacity	<input type="button" value="SELECT"/>
36197	+ Water birds	<input type="button" value="SELECT"/>
32369	+ Water blooms	<input type="button" value="SELECT"/>
28185	+ Water bombing	<input type="button" value="SELECT"/>

Relationship types	Date: 2006-06-13 17:28:56	
Language: <input type="text" value="English"/>	Term: Water activity (24929) - Descriptor with relations	
	Search: <input type="text"/>	<input type="button" value="Find Term"/>
Type of relationship: <input type="text" value="Broader Term (50)"/>		help
Language: <input type="text" value="English"/>	Related to: No term selected.	
	Search: <input type="text"/>	<input type="button" value="Find Related Term"/>

Water activity (EN) Word Tree See [Legend](#) for abbreviations

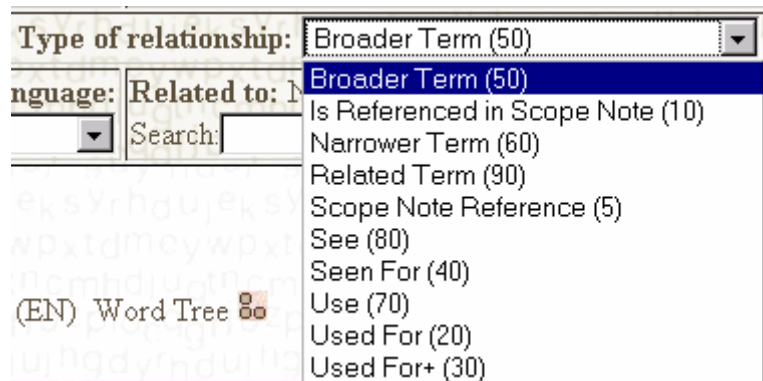
- Broader Term (50) 1998-11-05 09:10:00 -- [Chemicophysical properties \(EN\)](#)
- Related Term (90) 1998-11-05 09:10:00 -- [Moisture content \(EN\)](#)
- Related Term (90) 1998-11-05 09:10:00 -- [Water binding capacity \(EN\)](#)

All existing relationships are displayed.

For deleting a relationship click on the button. The system will delete the relationship and also its inverse.

It is not possible to modify an existing relationship here. Use the refinement tool instead, or delete the relationship and recreate the new one.

In order to create a new relationship you must first select the type of relationships from the scrolling list. The choice is between traditional thesaurus relationships:

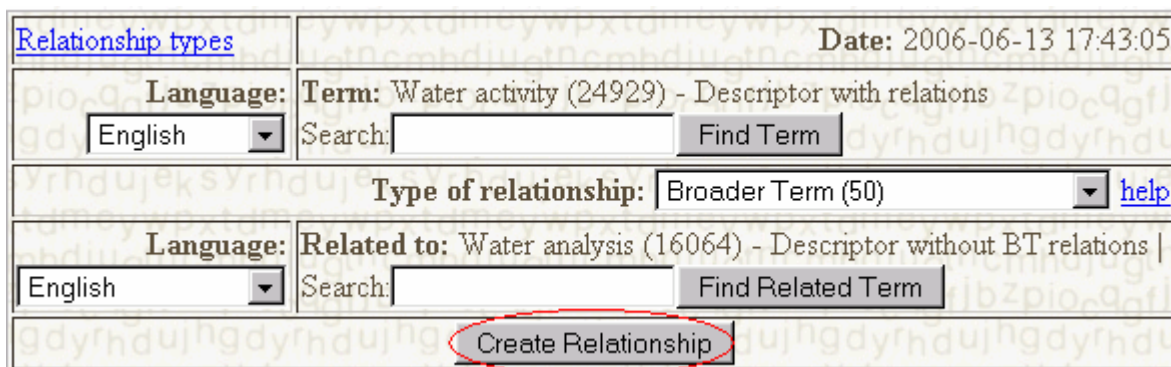


The definitions of these relationships can be seen by selecting the link “See Legend for abbreviations”.

After this choice you must select the term with which you want to create the relationship. Enter a term or part of a term in the “**Related to:**” field and press the “Find related term” button.

If no related term is selected the system shows “No term selected.”, and the relationship cannot be created.

When the “Type of relationship” and the “Related to” choices are done, click on the “Create Relationship” button.



The system will create the inverse relationship as well.

5.2. Manage relationships using the refinement tool


The refinement tool appears with a first page asking for a login and password:

AGROVOC REFINEMENT TOOL

Enter your login and password


User Login :

Password :



an application to browse the agrovoc database

AGROVOC ONLINE



an application to browse the agrovoc database

AGROVOC BROWSE



an application to browse the agrovoc database

AGROVOC MAINTENANCE TOOL

The refinement tool can be used to modify existing relationships. This information is stored in a separate field in the relationship table.

Use the same login and password as for the maintenance tool. After logging-in you see the main menu:

AGROVOC REFINEMENT

Welcome, GILW

1. [Relationship Refinement Tool](#)
2. Verification Tools
 - [By Term](#)
 - [By Semantic Relation](#)
 - [By Rule](#)

Select “Relationships refinement tool” and search for the term you want to refine. See picture below.

Relationship Refinement Tool

Search:

Criteria: Start with Any position

>> Back to Menu

- Water
- Water activity
- Water analysis
- Water availability
- Water balance
- Water binding capacity
- Water birds
- Water blooms
- Water bombing
- Water budget
- Water buffaloes
- Water cabbage
- Water chestnut
- Water chlorination
- Water circulation
- Water conservation

Water activity

Current Relation	New Relation	LEGEND	Related Term
BT	<input type="text" value="-----Select New Relationship-----"/>		Chemicophysical properties
RT	<input type="text" value="-----Select New Relationship-----"/>		Moisture content
RT	<input type="text" value="-----Select New Relationship-----"/>		Water binding capacity

Use the scrolling list to change existing relationships.

Water activity

Current Relation	New Relation	LEGEND	Related Term
BT	<input type="text" value="-----Select New Relationship-----"/>		Chemicophysical properties
RT	<input type="text" value="-----Select New Relationship-----"/>		Moisture content
RT	<input type="text" value="-----Select New Relationship-----"/> <div style="border: 1px solid black; padding: 5px;"> <p>-----Select New Relationship-----</p> <p>-----Hierarchical Relationships between Concept-----</p> <p>subclass of (/superclass of>)</p> <p>superclass of (/subclass of>)</p> <p>-----Concepts-to-Concepts Relationships-----</p> <p>affected_by (/affects>)</p> <p>affects (/affected_by>)</p> <p>afflicted by (/afflicts>)</p> <p>afflicts (/afflicted by>)</p> <p>beneficial for (/benefits from>)</p> <p>benefits from (/beneficial for>)</p> <p>caused by (/causes>)</p> <p>causes (/caused by>)</p> <p>component (/component of>)</p> <p>component of (/component>)</p> <p>compose (/composed of>)</p> <p>composed of (/compose>)</p> <p>derived from (/source>)</p> <p>goal or process (/means>)</p> <p>grows in (/growth environment for>)</p> </div>		Water binding capacity

6. AGROVOC structure and content revision

Over the years, different people with different backgrounds and knowledge have worked on AGROVOC. This has led to some inconsistencies in the use of the relationships or the term forms.

One of the main problems with multiple language thesauri is the need to be able to represent concepts which may not exist in all languages or may use different terms which are not the translation of each other. Also, the fact that different language versions of AGROVOC are generally carried out by translating the English version means that it is not completely and "purely" speaking a multilingual thesaurus, but a translated thesaurus.

For these reasons, we have been working for some years on the evolution from the thesaurus to what we call a "Concept Server" (CS), an ontology which makes clear distinction between concepts and the terms used to represent those concepts. Its real aim is to cover knowledge at three layers: concept, terms, and string, where a concept can be represented by multiple terms, terms can be in many languages and have synonyms, and a term can be represented with different spelling variants.

With the development of the FAO Agricultural Ontology Service⁴ (AOS), the aim is to build an infrastructure in which experts can add all the concepts they need, representing them in their native or local language, and have the ability to identify translations, if existing, of the terms used to represent concepts. Within the AOS initiative we plan to develop an infrastructure to allow distributed editing of the CS: specific registered and trained people (terminologist or ontology experts) will then be able to connect remotely and edit the CS directly. Specific "validators" can then just review the work and approve additions or changes.

Moreover, we have been starting activities for restructuring AGROVOC: the content of the thesaurus could be improved and inconsistencies could be avoided (singular and plurals for example, capitalizations, etc.). The relationships (BT, NT, RT, USE and UF) need also to be reviewed to be more consistently used. We are also currently revising the top terms: the next release of the thesaurus will be improved in its hierarchical structure.

There is also an ongoing effort to improve the semantics of AGROVOC. Today the thesaurus is basically term-oriented. Only one translation per term is possible and it is hardly possible to identify synonyms (multiple translations can be added as non-descriptors, but using the simple thesaurus relationships it may be difficult to identify clearly translations and synonyms, or spelling variants). Therefore, we are planning to restructure the thesaurus to have a more clear distinction between concepts, terms, and strings. The thesaurus should also reflect all different cultural perspectives and should not be driven only by the English view. The final result will be a concept-based repository, the CS as we call it, in which we can find clearly defined concept-to-concept, term-to-term and string-to-string relationships, or what we call lexical variances. Relationships between concepts will be made more explicit, in order to make better use of them (no more general "RT" relationships, but for example "hasPest", "madeOf", "ingredientOf"). For this, we also developed an extension of the current thesaurus maintenance tool that can be used to refine the current existing relationships in a more specific way. You can find information on the Web site (e.g. http://www.fao.org/aims/cs_relationships.htm), and you can find an article about that in the publication section.

⁴ Agricultural Ontology Service <http://www.fao.org/aims/aos.jsp>.

The maintenance of the CS will be distributed, done by "centres of excellence" through a specific system called the **AGROVOC Concept Server Workbench**. E-Conferences, training and workshops will be planned in order to nominate "AGROVOC managers" and promote quality development and maintenance of AGROVOC concepts and their structure. The new CS will make use of text mining of corpus, to help the work of experts for the population of the concept lists.

A **preliminary 0.1 Alpha Test version of the AGROVOC Concept Server Workbench** is now available online: <http://naist.cpe.ku.ac.th/agrovoc/>

The tool supports building and structuring multilingual ontology and terminology systems in the area of agriculture in a distributed environment. With that we are moving away from a centralized development of AGROVOC to a Web 2.0 inspired way of networked and distributed contributions to create a system with rich semantics that will greatly enhance indexing, searching, and information organization in the agricultural domain.

Please feel free to log in as guest and provide your valuable feedback. At this stage you can be part of the development process with your ideas and comments.

In order to enrich the CS and AOS in general, integration of other terminologies will be possible. Mapping to other thesauri, as we are doing with the Chinese Agricultural Thesaurus, ASFA Thesaurus and NAL Thesaurus, could also help to enrich the provided services.

From the CS it will not only be possible to export and further develop sub-domain specific ontologies but also to extract and access the "traditional" AGROVOC thesaurus and other terminology systems, like a simple term list, glossaries, etc., in SKOS, TBX or OWL format, as well as give access to the AOS resources through Web services.

Concerning the more semantic aspect of the CS, we are working to identify what would be the methodology and the process for converting a thesaurus into an ontology. Some tests and research in this area have already been carried out. Most of the information, specific presentations or articles can be found on our Web site: <http://www.fao.org/aims> under the Publications area. Currently we are finalizing an article describing the OWL model underpinning the CS structure⁵. Eventually, it will be necessary to identify how to distinguish between classes and instances, and how to modularize the CS. Scientific names of Plants and Animals may be taken from more authoritative sources such as Species 2000⁶.

7. AGROVOC online

We have developed a new portal to facilitate more coherence in the Agricultural area, where information about metadata standards and knowledge organization systems are disseminated. This portal incorporates all the functionalities and services for our member countries and users interested in Information Management Standards in Agriculture:
<http://www.fao.org/aims>.

⁵ FAO OWL Model: http://owl-workshop.man.ac.uk/acceptedPosition/submission_31.pdf.

⁶ Species 2000 website <http://www.sp2000.org/>

8. Accessing AGROVOC from your applications/systems

We have developed and we suggest using Web services for accessing AGROVOC in order to be automatically updated with all new modifications. Information regarding the available Web services can be found at: http://www.fao.org/aims/ag_Webservices.jsp.

9. AGROVOC and copyrights

The copyright of the localized version of AGROVOC remains with the institution that has performed the translation. The owner should take responsibilities of the maintenance of the translated release and share it also through FAO.