Results from a Study of the Implementation of Metadata Application Profiles in Agricultural Learning Repositories

Nikos Manouselis[1], Kostas Kastrantas[1], Gauri Salokhe[2], Jehad Najjar[3], Christian M. Stracke[4], Erik Duval[3]

[1]Informatics Laboratory, Agricultural University of Athens, 75 Iera Odos Str., 11855, Athens, Greece
{nikosm, kkastrad@aua.gr}
[2]Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla, 00153 Rome, Italy
gauri.salokhe@fao.org
[3]Computer Science Dept., K.U.Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium
{jehad.najjar, erik.duval}@cs.kuleuven.be
[4]Information Systems for Production and Operations Management, University of Duisburg-Essen, Campus Essen, Universitaetsstr. 9, D-45141 Essen, Germany
Christian.Stracke@icb.uni-due.de

In technology-enhanced learning, metadata interoperability has been identified as an important issue. It allows not only the exchange and preservation of crucial learning and teaching information (such as competency profiles, learning activities, and descriptions of learning resources), but also its future re-use among a large number of different systems and repositories. In the field of agricultural education and training, learning technologies' specifications and standards have not yet been widely adopted. Few initiatives have reported implementing them, and in most cases only to describe learning resources by using IEEE Learning Object Metadata (LOM), Dublin Core (DC) or a combination of the two. This results in introduction of dispersed efforts and systems with significant differences between them. In this paper, we present results of a study aiming at the review and assessment of implementations of metadata standards in agricultural learning repositories. More specifically, it assesses the current status of development and implementation of metadata application profiles based on standards such as IEEE LOM and DC.

***Index Terms*—Metadata, learning repository, application profile, agriculture.**

I. Introduction

Efforts in the application of technology enhanced learning methodologies and standards in the domain of agricultural education and training are, at present, widely dispersed. The amount of digitally available learning resources is growing exponentially. Appropriate steps need to be undertaken to facilitate sharing them, consequently reducing new content development efforts and costs.

These learning resources are often organized in databases that are called learning repositories (LRs). In the agricultural domain, similar applications often use dissimilar data models - which hamper information sharing. Additionally, information system designers are sometimes unaware of latest advances related to learning technology (LT) specifications and standards, which results in their slow adoption and implementation. As a matter of fact, few initiatives have reported implementing LT standards, more often to describe learning resources they use metadata standard such as IEEE Learning Object Metadata (LOM) [1] and Dublin Core (DC) [2].

In this direction, the Food and Agriculture Organization of the United Nations (FAO), has recently supported the launch of an Agricultural Learning Repositories Task Force (AgLR-TF, http://aglr.aua.gr). This community of experts, working on learning resources related to rural and agricultural populations, has been collecting, reviewing, and studying metadata application profiles (APs) implemented in agricultural learning repositories (AgLRs) around the world and trying to achieve some harmonization around their development.

In the context of these activities, an expert team has been also been assembled inside the Workshop on Learning Technologies (WS-LT) of the European Standardisation Committee CEN/ISSS. Its aim is to support AgLR-TF in the study of agricultural metadata APs, and thus help achieve the wider dissemination of LT standards for applications of the agricultural education and training domain. This team has carried out a study and assessment of existing implementations of metadata standards in AgLRs. This paper presents the first outcomes of this work.

More specifically, in this paper the current status of development and implementation of metadata application profiles (APs) for describing agricultural learning resources is carried out. A selection of representative APs, implemented in repositories around the world, has been studied in depth. Based on the results of this study, the paper makes an assessment of the compliance of the covered APs with their base schemas, and outlines a number of recommendations that can help adopters of LT standards in agricultural applications so that they design and implement better agricultural APs.

II. Rationale

A. Why Agricultural Learning Repositories

Making learning resources available online on a global scale could be an enabler for the development and welfare of agricultural and rural populations. One mean to achieve this aim, is the development and promotion of a global

infrastructure that will facilitate sharing and reusing of learning resources on topics related to agricultural and rural development.

From this perspective, the exchange of knowledge and experiences between the stakeholders involved in the development and operation of AgLRs is of high importance. These stakeholders can promote (in their own communities), and share (with other communities) guidelines, standards, technologies, tools, recommendations and good practices on a variety of topics:

- Designing, developing, adapting, and repurposing learning resources;
- Setting up learning repositories using open standards and technologies;
- Populating learning repositories with reusable learning resources and interoperable metadata;
- Designing and implementing a variety of quality assurance procedures and criteria for learning resources;
- Inter-connecting learning repositories in global federations to share and exchange resources and metadata;
- Deploying regional portals that provide access to learning resources in repositories around the world through federated services or metadata harvesting.

A number of agricultural learning repositories (AgLRs) have already been developed worldwide, as illustrated in Table I (source: http://aglr.aua.gr). The stakeholders involved in the development, operation and population of these repositories are getting more and more interested in the exchange of knowledge and experiences around the above listed topics of joint interest. A number of activities, such as an Agricultural Learning Repositories e-Conference (http://aglr.aua.gr/node/24) have already been organized in order to initiate the dialogue around these topics of interest.

TABLE I
OVERVIEW OF CURRENTLY OPERATING AGLRS

AgLR Name	URL	Country
Centre National de Recherche Agronomique (CNRA)	http://www.cnra.ci	Cote D'Ivoire
CGIAR On-line Learning Resources	http://learning.cgiar.org	United States
COTR's e-training site	http://kirk.estig.ipbeja.pt/cotr/	Portugal
EcoLearnIT	http://ecolearnit.ifas.ufl.edu	United States
FAO Capacity Building Portal	http://www.fao.org/capacitybuilding/	Italy
Lao Agriculture Database	http://lad.nafri.org.la	Lao People's Democratic Republic
Network of Aquaculture Centres in Asia-Pacific	http://www.enaca.org	Thailand
Rural-eGov Observatory	http://rural-egov.eu	Greece
SANREM CRSP Knowledge Base	http://www.oired.vt.edu/sanremcrsp/menu information/SKB.php	United States
TrAgLor - Turkish Agricultural Learning Object Repository	http://traglor.cu.edu.tr	Turkey

B. *Context of this work*

LT specifications and standards could help achieving some level of interoperability and harmonization in the various AgLR implementation efforts. On the other hand, LT have not yet been widely adopted and implemented in the agricultural education and training domain. Most of the activities that have reported implementing them are only mostly using metadata standard such as IEEE LOM [1] and DC [2] to describe learning resources. The AgLR-TF community of experts has been collecting and studying implemented AgLRs, particularly focusing on the metadata APs that they use.

In order to assess, validate and contribute to the work that the AgLR-TF community is doing, a parallel expert group has been formulated in the context of the CEN/ISSS WS-LT. More specifically, WS-LT decided in February 2008 to set up a liaison with the AgLR-TF community of experts, in order to further promote the adoption and implementation of LT specifications and standards in learning repositories that support the needs of rural and agricultural populations. Based on the work carried out in the CEN Workshop Agreement 15555 “Guidelines and support for building application profiles in e-learning” [3], it was agreed to set up a joint project team that will review implemented metadata APs for a sample of popular AgLRs, in the light of the guidelines of CWA 15555. Then, it would try to elaborate a set of recommendations for achieving better interoperability between them. This paper presents the first outcomes of this activity.

III. METHODOLOGY

This section provides an overview of the methodology that has been followed for the analysis of the sample of agricultural APs.

Based on CWA 15555 and [4], the project team elaborated a number of analysis dimensions that have been incorporated into an appropriate analysis tool. This aimed at supporting the analysis of the agricultural APs, through a template that has been developed as an Excel file. It included the following components:

- An overall overview of the analysed AP, which includes general information (such as its title, description, and producer), information about existing documentation (such as a conceptual model and data bindings), information about its scope and purpose (such as a clear scope definition and use cases), and an overview of the results of the mapping of the AP into its base schema (particularly focusing on allowed and non-allowed modifications).
- A detailed mapping of the analysed AP onto its base schema(s), i.e. IEEE LOM, DC, or both.

Each AP in the sample has been assigned to at least one member of the project team. An analysis report (including both general recommendations as well as a mapping of the AP to its base schema) has been produced. Finally, all AP reports have been integrated into one overall report.

The liaison with AgLR-TF was the source of the sample of agricultural APs evaluated as follows:

1. Rural-eGov IEEE LOM AP (ReGov LOM).
2. FAO Agricultural Learning Resources AP (FAO Ag-LR).
3. CGIAR LOM Core AP (CG LOM Core).
4. BIOAGRO LOM AP.
5. Biosci Education Network (BEN) AP.
6. Sustainable Agriculture and Natural Resource Management Collaborative Research Support Program (SANREM CRSP) AP.
7. TrAgLor LOM AP.
8. Intute: Health and Life Sciences AP (Intute AP).
9. EcoLearnIT LOM AP.

From those mentioned above, it was possible to analyse in detail only the first six. For the TrAgLor LOM AP only a preliminary analysis took place, based on basic information that was provided by the AP developers (such as a database instance of its implementation). For the Intute and EcoLearnIT LOM APs, no analysis was possible (they are included here for reference reasons).

IV. Analysis of Agricultural APs

This section provides the results from the analysis of the APs of the collected sample. For each AP the corresponding information of the overall component of the analysis report is presented. Corresponding mappings have also been elaborated (not presented in this paper).

A. ReGov LOM

Rural-eGov IEEE LOM (or simply ReGov LOM) AP is an IEEE LOM-based AP that has been developed to facilitate the description and categorization of learning resources that have been developed to support the training of rural and agricultural small and medium enterprises (SMEs) on topics related to the use of e-government [5]. Resources are being developed for the SMEs in five rural areas of Europe: UK, Greece, Poland, Slovenia and Germany. They are being populated in a repository of learning resources, called the Rural-eGov Observatory. In addition, they are also described with metadata in one of the five targeted languages (i.e. English, Greek, Polish, Slovenian and German).

The application profile has been published on 25/9/2007 by the Informatics Laboratory of the Agricultural University of Athens, in the context of the project "Rural-eGov: Training SMEs of Rural Areas on Using e-Government Services" of the Leonardo da Vinci Programme (http://rural-egov.eu).

No technical binding is provided for ReGov LOM, although it might provide helpful for other developers. The implementation of ReGov LOM is in a relational database, therefore no technical conformance is claimed for the XML binding of LOM. Apart from conceptual data model, its documentation seems to provide a clear scope of its purpose. Furthermore, a detailed specification document of the Rural-eGov Observatory system provides a number of use cases that helped in eliciting some additional requirements for the development of the AP. The only limitation is that the introduction of the use cases is not in the document that describes the AP itself – this is a shortcoming that could be addressed by the developers of the schema.

The ReGov LOM AP contains only elements from the IEEE LOM standard. In specific, from a total of 77 elements, ReGov LOM uses 65 elements of LOM. Based on this selection, ReGov LOM can be considered as a subset AP of IEEE LOM. Contrary to the IEEE LOM Standard where the use of all of its elements is optional, ReGov LOM defines 39 mandatory elements, such as *1:General, 5:Educational* etc. Additionally ReGov LOM reduces the size of some elements to one, contrary to LOM that defines their size as multiple, e.g. *1.1:General.Identifier, 1.4:General.Description*, etc.

On the other hand, the value space of some elements is restricted to a subset of the values that are provided by the IEEE LOM for the given elements, e.g. value space of *2.3.1:LifeCycle.Contribute.Role* element is restricted to "author", "publisher", "unknown", "validator", "editor", "subject matter expert", "annotator". Additionally other elements' value spaces are defined as references to another standard, e.g. value space of *1.5:General.Keyword* is defined from the AGRIS Subject Categories (http://www.fao.org/scripts/agris/c-categ.htm) of FAO. For this reason, the value space of ReGov LOM can be considered as multi-source. Overall, all these modifications conform to the CEN/ISSS CWA 15555 guidelines regarding the modification of the value spaces.

As far as element datatypes are concerned, ReGov LOM has changed the datatype of *1.5:General.Keyword*, *1.6:General.Coverage*, *9.2.2.1:Classification.TaxonPath.Taxon.Id* and *9.2.2.1:Classification.TaxonPath.Taxon.Entry* from 'LangString' to 'Vocabulary'. This is a non-allowed change, since the modification of datatypes is not conformant to the CWA 15555 guidelines. It could be suggested to the developers of the AP to keep the datatype as 'Langstring', and that they allow only the desired text values in the metadata input system interface.

B. FAO Ag-LR

Capacity and institution building is a core function of FAO. FAO has recently launched the Capacity and Institution Building Portal (http://www.fao.org/capacitybuilding/) to provide structured access to information on FAO's capacity and institution building services and learning resources. To ensure that the Portal can be searched by users and to enable interoperability with other recognized educational repositories, the FAO Agricultural Learning Resources (Ag-LR) AP was created conforming to available and commonly used standards, to describe agricultural learning resources [6].

The Ag-LR AP of FAO was published in September 2007 and aims to serve as an international reference for designing and developing repositories of agricultural learning resources. Its schema is based mainly on the Dublin Core Metadata Element Set (DCMES) and the Agricultural Metadata Element Set (AgMES), with additional elements taken from the IEEE

LOM standard. Although an XML technical binding has been developed for internal use, it is not yet made public for other potential implementers. The binding claims conformance to the DCMES base schema.

Within the FAO Ag-LR AP documentation, clear statements about its purpose are provided, along with a set of basic requirements that aim to cover, as well as a set of examples of its use within FAO's portal. On the other hand additional requirements regarding the use of FAO Ag-LR's elements could be presented through exemplary use cases.

FAO Ag-LR AP is made up of 22 elements. As stated above it is derived by combining elements from DCMES and AgMES. Additional elements from the IEEE LOM standard have also been selected, mainly to capture the educational characteristics of the agricultural learning resources. Since its elements are selected from more than one standard, the FAO Ag-LR application profile can be considered as a multi source application profile. The same also applies for its value space, since its values come from various sources (such as the LOM or FAO namespaces). The process of selecting the elements and value spaces of the FAO Ag-LR AP is conformant to the CWA 15555 guidelines.

The AP defines 9 mandatory elements, restricting the requirement of the base schemas that define all their elements as optional. In addition, it does not break any super-class/sub-class semantic relations.

Modifications to the value spaces of the elements are also conformant to the CWA 15555 guidelines, since they were defined as a reference to existing specifications or as a subset of the defined value space of the base schema. For example the values of element *4.Subject/Categories* are defined from FAO's Technical Knowledge Classification Scheme, while the values of *19.IntendedEndUserRole* are restricted to "teacher", "learner", and "manager".

C. CG LOM Core

The Consultative Group for International Agricultural Research (CGIAR, http://www.cgiar.org) is a strategic alliance of countries, international and regional organizations, and private foundations supporting 15 international agricultural research centers in developing countries. CGIAR has initiated the Online Learning Resources (OLR) project which addresses the need of an international teaching and learning community of practice, interested in tropical agriculture and natural resources management research and development, to easily discover and retrieve relevant learning resources produced by the various centers in collaboration with their national partners.

In the context of the OLR project, CGIAR published the CG LOM Core AP to enable the description of the CGIAR learning resources for populating the CGIAR learning object repository . CG LOM Core is developed upon the IEEE LOM standard and was published on November 2005 [7]. It generally tries to cover the needs of the CGIAR teaching and learning community of practice.

Within its documentation, the conceptual data model of the CG LOM Core AP is being thoroughly presented. On the other hand no technical binding is publicly available, although the developers of the schema claim technical conformance to the base schema. A clear purpose statement of the CG LOM Core is being provided within its documentation. On the other hand, no use cases are included to capture/illustrate the specific needs of the targeted community.

CG LOM Core's elements are directly derived from the IEEE LOM standard. In specific, from a total of 77 elements, CG LOM Core uses 74 elements of IEEE LOM standard. For this reason, CG LOM Core can be considered as a subset AP of IEEE LOM. Contrary to the IEEE LOM standard where the use of all of its elements is optional, CG LOM Core defines 36 mandatory elements. On the other hand the value space of the AP can be overall considered as multi-source. On one hand, the majority of its elements use the value space as it has been defined by the IEEE LOM standard. On the other hand, one element, namely *7:Relation.Kind* extends the predefined value space with additional two elements and another one, *9:Classification.TaxonPath.Taxon* uses the values of AGRIS Subject Categories as its value space.

The analysis showed that while the obligation status of *2.3.1:LifeCycle.Contribute.Role, 2.3.1:Life Cycle.Contribute. Entity* and *2.3.1:LifeCycle.Contribute.Date* has been modified to mandatory their parent element *2.3:LifeCycle.Contribute* retained its status as optional. This is not conformant to the CWA 15555 guidelines and it should be changed to mandatory. Moreover, the size of the elements *5.9:Educational.TypicalLearningTime* and *5.10:Educational. Description* has been modified from one to multiple. Again, this is a modification that is not conforming to the CWA 15555 guidelines.

In addition, the modification of datatypes from 'Langstring' to 'Vocabulary' for elements *5.9:Educational.TypicalLearningTime* and *5.10:Educational. Description* is also a non-allowed modification. It could be suggested to the developers of the AP to keep the datatype as 'Langstring', and to provide as a recommendation to the implementers that they allow only the desired text values in the metadata input system interface. Finally, as stated above the addition of two more values in the value space of element *7:Relation.Kind* is again a non allowed modification since they do not derive from a standard value space. It is suggested that the developers of the schema make their new value spaces available in a public namespace, in order for other implementations to be able to find, reference, and use them.

D. BIOAGRO LOM

BIOAGRO LOM is an IEEE LOM-based Application Profile that has been developed in the context of the Bio@gro eContent project, to facilitate the annotation/description of learning resources that are collected and described in the Bio@gro Web Portal . The BIOAGRO LOM AP has been particularly developed to support the description of learning resources on the topic of organic agriculture. Multilingual descriptions in four languages (i.e. English, German, Greek

and Romanian) are stored in the portal.

The application profile has been published on 20/12/2005 by the Informatics Laboratory of the Agricultural University of Athens [8]. Within its documentation the conceptual data model of the AP is been thoroughly presented. On the other hand no technical binding of the application profile is being provided. The BIOAGRO LOM AP documentation seems to describe clearly the scope of its purpose. However, it does not include any use cases.

The BIOAGRO LOM AP contains only elements from the IEEE LOM Standard. In specific, from a total of 77 elements, BIOAGRO LOM uses 35 elements of LOM. For this reason, BIOAGRO LOM can be considered as a subset AP of IEEE LOM.

On the other hand, the value set of this AP can be considered as ad-hoc since the value space of several elements (such as *1.5:General.Keyword*, *5.5:Educational. IntendedEndUserRole*) has been created to serve the specific requirements of the project. However the rest of the elements (with two exceptions that are discussed below) follow the complete value space as it has been defined by the IEEE LOM.

BIOAGRO LOM defines 16 mandatory elements. Additionally it reduces the size of some elements to one, contrary to IEEE LOM that defines their size as multiple, e.g. *1.4:General.Description, 2.3:Life Cycle.Contribute*, etc.

BIOAGRO LOM has made non allowed data type modifications from 'Langstring' to 'Characterstring' in elements *1.2:General.Title*, *1.4:General.Description*, *4.6:Technical.OtherPlatformRequirements*, *5.6:Educational. TypicalAgeRange* and *6:Rights.Description*). It would be an option for the developers to make sure that when they extract content and/or exchange metadata with other repositories, they make sure that the stored values for these elements are transformed into 'Langstring' datatypes.

This AP also changed the datatype of other elements from 'Langstring' to 'Vocabulary'. It could be suggested to the developers of the AP to keep the datatype as 'Langstring', and to provide as a recommendation to the implementers that they allow only the desired text values in the metadata input system interface.

Additionally, the value space of elements *3.2.1:Meta-Metadata.Contribute.Role* and *5.2:Educational. LearningResourceType* (as been defined by IEEE LOM) has been extended with additional values to meet the needs of the specific project. However this modification does not conform to the CWA 15555 guidelines since these values are not referenced from any other known standard or specification nor they were published in a public namespace. It is suggested that the developers of the schema make their new value spaces available in a public namespace, in order for other implementations to be able to find, reference, and use them.

Finally, the BIOAGRO LOM AP has modified the cardinality of 18 elements of the base schema. However the cardinality modification of 3 elements was not in line with the CWA 15555 guidelines. In specific, contrary to the base schema, the cardinality of elements *2.3.1:LifeCycle.Contribute.Role*, *4.2:Technical.Size* and *4.6:Technical.OtherPlatformRequirements* has been extended to a size larger than one. This is a non-allowed modification that has to be carefully considered when extracting content and/or exchanging metadata with other repositories. Inevitably, information will be lost during such a transformation/mapping.

E. BEN

The Biosci Education Network (BEN) Collaborative established a portal site and is developing and maintaining digital library collections of biological sciences teaching and learning resources. Part of the development of the portal and digital library collections entailed the establishment of a metadata specification (the BEN AP) that all partners agreed to adhere to for describing their collection's material. BEN AP aims to support the cataloging of BEN learning resources and user discovery (searching and browsing) of biology teaching and learning resources.

BEN application profile was published on April 2003 and it is based on a previous version of IEEE LOM Standard: more precisely, on April's 2001 version 6.1 a direct precedent version of the current standard. Their difference lies in the naming of some elements however their semantics remain the same. For example, the *Identifier* element is called *CatalogEntry* but they both have the same semantics.

The documentation describes a clear scope and purpose of the AP and its conceptual data model is being thoroughly presented. No technical binding of the conceptual data model is provided.

Although the document states that ad hoc specialized working groups were teamed up to discuss and conclude on the conceptual data model, details on how user requirements were collected along with some indicative use cases could also be provided.

All available elements from the IEEE LOM v6.1 metadata schema have been selected for the BEN AP, and for this reason it can be considered as a complete AP. However, it cannot be considered as a conforming IEEE LOM AP, due to structural differences with the current version of the standard.

On the other hand, regarding the value space of the BEN application profile, in some occasions ad hoc vocabularies had to be developed in order to cover the particular requirements of BEN's user groups, e.g. *5.2:Educational.LearningResourceType*, *5.6:Educational.Context*, *9.1:Classification.Purpose*. For the rest of the elements the complete value space of IEEE LOM's v6.1 is being used.

Contrary to the IEEE LOM v.6.1 schema where the use of all of its elements is optional, BEN AP defines 31 mandatory elements. In addition, some modifications to the value spaces of elements have taken place. For instance, the value space of *2.2:LifeCycle.Status* has been modified to the one used by the Digital Library for Earth System Education (DLESE) metadata specification. In *2.3.1:Contribute.Role*, BEN uses a

subset of the LOM-proposed vocabulary.

In two elements, ad hoc value spaces have been defined in order to meet the needs of the user community. More specifically, these have been *5.2:Educational.Learning* and *5.6.Educational.Context*. We could not locate their value spaces in some public namespace. Therefore, it would be suggested to the developers of the schema to make their own value spaces available in a public namespace, in order for other implementations to be able to find, reference, and use them.

F. SANREM CRSP

The Sustainable Agriculture and Natural Resource Management Collaborative Research Support Program (SANREM CRSP) is sponsored by the U.S. Agency for International Development's Economic Growth, Agriculture, and Trade Bureau (USAID/EGAT) and participating U.S. and host country institutions around the world. The objective of the SANREM CRSP is to support sustainable agriculture and natural resource management decision makers in developing countries by providing access to appropriate data, knowledge, tools, and methods of analysis; and by enhancing their capacity to make better decisions to improve livelihoods and the sustainability of natural resources.

All SANREM programs and activities contribute to the online SANREM Knowledgebase (SKB, http://www.oired.vt.edu/sanremcrsp/menu_information/SKB.php), which it is intended to serve as a catalog of information resources specific to the SANREM project as well as catalog and archive other resources and projects that relate to sustainable agriculture and natural resource management. The "resources" cataloged in the SKB (using DC-based metadata) are primarily articles, papers, and reports but may include other digital resources such as presentations, images, web pages, and other materials that can be referenced. The overall goal is to make these resources readily available to facilitate the wide and effective dissemination of information and to provide a structure for effective search and retrieval of the resources. The SANREM AP is defined in the SANREM Knowledgebase Metadata Guide (http://www.oired.vt.edu/sanremcrsp/documents/SKB.metadata.guide.V4.Oct.2007.pdf) that has been published in September 2007.

Within its documentation the conceptual data model of the SANREM KB application profile is being presented. Additionally, conformance is claimed to the DCMES-however for sake of interoperability, a more in depth analysis of the application profile could be presented following the CWA 14855 "Dublin Core Application Profile Guidelines" [9]. On the other hand, no technical binding of the conceptual data is provided, nor conformance is claimed to the technical binding of the base schema. A clear purpose statement of the application profile is being provided. However, no use cases are being provided for capturing the requirements of the targeted user groups.

As mentioned above, the SANREM KB application profile is conformant to the simple DCMES. Additionally, some DC qualifiers have also been included (e.g elements Spatial and Temporal as refinements of DC element Coverage, and element IsPartOf as refinement of DC element Relation. Additionally some ad-hoc elements have been introduced for serving the specific needs of the SANREM community, such as the elements SANREMProductType, SANREMProjectID, UploadResource etc. Overall, the SANREM KB application profile can be considered as a mixture of a subset AP and an ad-hoc AP of IEEE LOM.

The SANREM KB Application Profile defines 6 mandatory elements, namely Title, Description, Creator (author), CreationDate, Description and Type. No additional modifications have been made to the value space of the base schema.

A vocabulary set has been defined for describing the keywords of a given resource. However this modification does not conform to the CWA 15555 guidelines since this value space is not referenced by any other known standard or specification nor they were published in a public namespace. It is suggested that the developers of the schema make their new value spaces available in a public namespace, in order for other implementations to be able to find, reference, and use them.

G. TrAgLor

The Turkish Agricultural Learning Objects Repository (TrAgLor, http://traglor.cu.edu.tr/) is a Turkish initiative project, funded by the Scientific and Technological Research Council of Turkey. It is coordinated by Agricultural Faculty of Çukurova University. In collaboration with several educational, private and public sector organizations, it aims to promote an infrastructure for learning objects in agriculture, food, environment, forestry and veterinary sciences.

For the storage and the description of the associated learning objects in the repository, the TrAgLor LOM AP has been developed [10]. As with the BEN Application Profile, TrAgLor is based on April's 2001 version 6.1 of IEEE LOM Standard.

The TrAgLor project team has published some research papers that present a detailed architectural analysis of the TrAgLor project. However, there is no formal and sufficient documentation describing the TrAgLor LOM AP itself so far. On the other hand, the XML binding of a TrAgLor LOM AP metadata instance is available from TrAgLor. Although TrAgLor-relevant papers describe clearly the scope and purpose of the repository, there is no documentation available describing the scope and purpose of the metadata AP itself.

All available elements from the IEEE LOM v6.1 metadata schema have been selected for the TrAgLor AP, and for this reason it can be considered as complete. However, it cannot be considered as a conforming IEEE LOM AP, due to structural differences with the current version of the standard.

Due to the lack of sufficient documentation, it is not possible to draw any conclusions about the value spaces, the allowed modifications, and the non-allowed modifications of

the base schema elements.

H. Intute: Health and Life Sciences

Intute (http://www.intute.ac.uk/) is a free Web-enabled service aimed at students, teachers, and researchers in UK further education and higher education. Intute provides to online access to a large database of resources that cover four main subjects: (a) Science and Technology (b) Arts and Humanities (c) Social Sciences and (d) Health and Life Sciences. One particular collection of the last subject is the one listing resources on Agriculture, Food and Forestry – formerly known as the AgriFor service (http://www.intute.ac.uk/healthandlifesciences/agriculture/).
In early June 2008 Intute provided 123,381 records. Intute is not to be confused with a simple search engine, since subject experts continuously select and include resources describing them with appropriate metadata. Agriculture-related annotation also takes place, for instance according to the CABI Thesaurus. The Intute AP is based on DC. Due to insufficient documentation, a further analysis of the AP was not possible.

I. EcoLearnIT

EcoLearnIT (http://EcoLearnIT.ifas.ufl.edu) is a digital repository of reusable learning objects that manages and hosts various resources focused on soil, water and environmental sciences, and provides authoring tools to develop learning objects. It is an open-access system in which learning objects are created, reviewed and used by an international community of online learners, students, instructors and scientists. EcoLearnIT facilitates learning at all levels ranging from simple to complex knowledge encapsulated into different types of resources, targeting various learning audiences (ranging from post-graduate, graduate and undergraduate students, online learners, farmers, ecologists, consultants, agencies and others). EcoLearnIT is implemented in form of an online journal that provides full credit to authors and co-authors (i.e., the learning object developer). The resources stored in EcoLearnIT can be cited and referenced (hyperlinked) to be included in courses, training and extension material and accessed by students, learners and instructors. The EcoLearni IT metadata AP is based on IEEE LOM. Due to insufficient documentation, a further analysis of the AP was not possible.

V. Outcomes

This section discusses the overall observations of the previous analysis, and integrates the suggestions upon each AP into a set of generic recommendations that could be useful for the designers and developers of metadata APs for AgLRs.

A. General Observations

The analysis of the sample of studied APs has led to a number of interesting observations, as far as the CWA 15555 guidelines are concerned.

Documentation. The majority of the nine (9) identified APs for AgLRs had accompanying documentation that described/specified the AP. In three (3) cases, there was no documentation. After communicating with the operators of the repositories, we have been informed that for TrAgLor and EcoLearnIT, they intended to prepare and publish such documentation in the near future. As far as Intute is concerned, there was no plan for the preparation of a specific document describing the metadata AP, since it is based on the more generic ePrints metadata schema (which is a DC AP).

For the seven (7) APs of the sample that have been analysed (including TrAgLor to the extend this was possible):

- All of them had a description or specification of their conceptual model, as requested by CWA 15555;
- Only two (2) of them (i.e. FAO's Ag-LR, TrAgLor LOM) provide public access to a technical binding (in both cases in XML).
- Three (3) of them claim technical conformance to their base schema. That is, FAO's Ag-LR to DC; CG LOM Core and TrAgLor LOM to IEEE LOM.

Scope and Purpose. In all seven (7) APs there is a description of the scope and purpose, as suggested by CWA 15555. On the other hand, in only two (2) is there a set of use cases further elaborating the needs and requirements of the user community (i.e. ReGov LOM, FAO's Ag-LR).

Number of elements. As far as the selected elements from the base schemas are concerned, the analysed APs can be classified as (Najjar et al., 2004):

- Complete APs, three (3) of them. These are SANREM KB, BEN LOM and TrAgLor LOM.
- Subset APs, three (3) of them. These are ReGov LOM, CG LOM Core, and BIOAGRO LOM.
- Multi-source APs, one (1) of them. This is FAO's Ag-LR one.

It was interesting to observe that in five (5) of the APs, the considered elements where either all the LOM ones (i.e. SANREM KB, BEN, TrAgLor) or the majority of them (i.e. ReGov LOM, CG LOM Core). These included more than 65 elements (out of the 77 that LOM has).

On the other hand, the other two (2) APs proposed the use of fewer elements. BIOAGRO LOM includes thirty five (35) elements, whereas FAO's Ag-LR only twenty two (22).

Nevertheless, these numbers are indicative since the mandated use of elements is not related to the potential number of elements that an AP includes. In the examined sample:

- Three (3) APs require more than thirty (30) mandatory elements: ReGov LOM (39), CG LOM Core (36), BEN LOM (31).
- One (1) AP requires more than fifteen (15) mandatory elements: BIOAGRO LOM (16).
- Two (2) APs require less than ten (10) mandatory elements:

FAO Ag-LR (9), SANREM KB (6).

For TrAgLor, no information about mandatory elements is available.

Allowed Modifications from Base Schema. The allowed modifications from the base schema that have been observed most often are:

- The mandatory selection of non-mandatory elements: in six (6) APs.
- The change in the definition of the size and smallest permitted maximum: in two (2) APs.
- The change in the obligation of data elements: in six (6) APs.
- The allowed modifications of value spaces: in five (5) APs.
- Other modifications: in one (1) AP.

Non-Allowed Modifications from Base Schema. The non-allowed modifications from the base schema that have been observed most often are:

- The extension of the cardinality of elements: in one (1) AP.
- The addition of new items in controlled vocabulary lists: in two (2) APs.
- The change in the mandatory status of an element: in one (1) AP.
- The use of a base schema version other than the latest, stable one: in two (2) APs.

Apart from those, the rest of non-allowed modifications that CWA 15555 mentions have not been noted. That is:

- The change in the location of a data element.
- The creation of new elements that mimic the semantics of existing ones.
- The change in the meaning of existing elements.
- The change in the name of existing elements.
- The extension of a base schema in other than specified points.

B. Usage of Elements

Based on the analysis of which elements are more often used in the examined sample, we can make some initial observations about the elements that seem to be more popular in agricultural APs:

- Most of the APs are using some element to store an identifier of the resource. In some cases, this is only a URL (in other cases, a formal catalog system can also been used).
- As far as the rest of the general characteristics of the resource are concerned, the following information is usually stored:
  - Title;
  - Language;
  - Description;
  - Keyword (free text or restricted);
  - Coverage (geographical/spatial or temporal).
- As far as the life cycle of the resource is concerned, the following information is usually stored:
  - Role of the entities that have contributed to the resource;
  - Information about these entities;
  - Date of contribution/production/publication.
- As far as the technical characteristics of the resource are concerned, the following information is usually stored:
  - Technical format;
  - Technical location (such as URL), when the Identifier element is not used for this purpose;
  - Size;
  - Some technical requirements for its viewing/execution.
- As far as the educational characteristics of the resource are concerned, the following information is usually stored:
  - Type of the learning resource;
  - Intended end user role;
  - Educational context/level.
- As far as the copyrights of the resource are concerned, the following information is usually stored:
  - Cost;
  - Copyrights and restrictions in use
- As far as the formal classification of the resource is concerned, the following information is usually stored:
  - Purpose of classification;
  - The classification system used;
  - Terms used from the selected classification system.

C. Initial Suggestions

Based on the analysis of the sample of APs and the overall observations made above, we could come up with the following suggestions/recommendations to the designers/developers of metadata APs for AgLRs:

1. *Always provide supportive documentation describing the AP.* Supportive documentations offer and allow an overview for the selection and reference for detailed analysis within the adoption phase.
2. *Include in documentation reference to the technical implementation of the AP and provide any relevant technical bindings.* References to technical implementations and provided technical bindings facilitate the implementation and the technical interoperability.
3. *Include in documentation supportive use cases that help clarify its scope, purpose and users.* Use cases support the selection process during the comparison of AP candidates and provide information about implementation potentials.
4. *Use the latest and more stable version of the base schema available.* Different versions of metadata specifications and/or standards often have important differences that do not ensure backwards compatibility. When starting an implementation project, it is suggested that AP designers/implementers chose the latest and more stable version of the base schema that is publicly available. For instance, in one examined case, although the project was initiated after the publication of the IEEE LOM standard in 2002, a previous version of LOM has been used.
5. *When ad hoc or extended value spaces are used for some*

elements, it is required to make the new value spaces available in a public namespace, in order for conformance to be maintained. Public availability is needed to ensure interoperability of future APs by allowing references to these published namespaces.

6. *Instead of substituting the 'Langstring' datatype with the simpler 'Characterstring', it is suggested that simplicity is sought though appropriate interface design.* For instance, when the type is changed from 'Langstring' to 'Characterstring', then implementers have to make sure that during a transformation/mapping the stored values for these elements are transformed into 'Langstring' datatypes in order to avoid information loss.
7. *The non-allowed modification rules of CWA 15555 should be carefully respected, because breaking them can lead to problems when trying to export/exchange metadata.* For instance, an extension to the cardinality of an element can lead to loss of information during a transformation/mapping.
8. *The elements most often occurring as mandatory in the existing APs should be considered for use also in other APs, to facilitate information exchange and interoperability.* It is most probable that the information that is considered important in all other agricultural APs will also be important for a new one as well. To achieve interoperability in metadata exchange, information about a characteristic that is stored in all other APs will have to be stored for a new AP as well.

VI. Conclusions

In this paper, we present results of a study that aimed at evaluating existing implementations of metadata standards in AgLRs. More specifically, it assesses the current status of development and implementation of metadata standards and specifications (such as IEEE LOM and DC), in the case of describing agricultural learning resources. It studies in depth a selection of representative APs that have been implemented in repositories around the world. Then, it reports on the compliance of the developed APs with their base schemas, and makes some initial recommendations regarding the better design and implementation of such APs.

Acknowledgment

The authors would like to thank the experts of the CEN/ISSS WS LT for their valuable comments and feedback. Part of the work presented in this paper has been funded by the Food and Agriculture Organization of the United Nations (FAO).

References

[1] IEEE LOM, *Draft Standard for Learning Object Metadata*, IEEE Learning Technology Standards Committee, IEEE 1484.12.1-2002, 15 July 2002.

[2] DC, *Dublin Core Metadata Element Set, Version 1.1: Reference Description*, Dublin Core Org., 2004.

[3] CWA 15555 "Guidelines and support for building application profiles in e-learning", European Committee for Standardization (CEN), June 2006 (http://www.cen-isss-wslt.din.de/sixcms_upload/media/3050/CWA15555.pdf).

[4] J. Najjar, S. Ternier, and E. Duval, "Interoperability of Learning Object Repositories: Complications and Guidelines", *IADIS International Journal of WWW/Internet*, 2004 (http://www.cs.kuleuven.ac.be/~najjar/papers/IadiswwwJournal.pdf).

[5] N. Manouselis, K. Kastrantas, and A. Tzikopoulos, "An IEEE LOM application profile to describe training resources for agricultural & rural SMEs", in the Proc. of *the 2nd International Conference on Metadata and Semantics Research (MTSR'07)*, Corfu, Greece, October 2007 (http://infolab-dev.aua.gr/files/publications/en/1200946986.pdf).

[6] H. Stuempel, G. Salokhe, A. Aubert, J. Keizer, A. Nadeau, S. Katz, and S. Rudgard, "Metadata Application Profile for Agricultural Learning Resources", in the Proc. of the *2nd International Conference on Metadata and Semantics Research (MTSR'07)*, Corfu, Greece, October 2007 (ftp://ftp.fao.org/docrep/fao/010/ai154e/ai154e00.pdf).

[7] T. Zschocke, C. Paisley, E. Duval, and J. Beniest, *CG Learning Object Metadata (LOM) Core*, CGIAR ICT/KM OLR Project, 2005.

[8] Bio@gro, *Metadata Models for Bio@gro Content Objects (BCOs) Description*, Bio@gro Technical Report, 2005.

[9] CWA 14855 "Dublin Core Application Profile Guidelines", European Committee for Standardization (CEN), November 2003 (ftp://cenftp1.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa14855-00-2003-Nov.pdf).

[10] Z. Cebeci, Y. Erdogan, and M. Kara, "TrAgLor: A LOM-Based Digital Learning Objects Repository for Agriculture", in Proc. of the *4th International Scientific Conference on "eLearning and Software for Education" (eLSE'08)*, Bucharest, Romania, 17-18 April 2008. (http://traglor.cu.edu.tr/objects/pdf/Traglor-else08.pdf).