Hungarian Association of Agricultural Informatics
European Federation for Information Technology in Agriculture, Food and the Environment

Journal of Agricultural Informatics. Vol. 8, No. 3
journal.magisz.org

Knowledge databank and repository service for agroforestry

Charles Burriel[1], Miklós Herdon[2], János Tamás[3], László Várallyai[4]

I N F O
Received 17 Jul 2017
Accepted 10 Oct 2017
Available on-line 21 Nov 2017
Responsible Editor: M. Herdon

Keywords:
knowledge, databank, agroforestry, repository, thesaurus, taxonomy

A B S T R A C T

The agricultural system experienced a strong abandonment of agroforestry in the 20th century, with agroforestry today taking place only on a few million ha in Europe. One of the aims of the AgroFE (Agroforestry Training in Europe) and the Agrof-MM (Agroforestry – Training – Mediterranean and Mountain Erasmus+) European projects is to build a knowledge databank to assist educational development and to provide experts and farmers with agroforestry knowledge. The KDB (Knowledge DataBank) is based on different professional vocabulary, metadata and thesaurus systems to build the content structure and help the users in their searches. A new tool will improve the knowledge management of agroforestry. The Agrof-MM project is creating an agroforestry thesaurus. The paper describes the knowledge databank system prototype for Agroforestry training and education. The knowledge databank is a component of the project training system. It aims to gather and share a set of documents and resources that partners, learners and public users can access in different forms: mono document objects and composite materials.

1. Introduction

The agricultural system experienced a strong abandonment of agroforestry in the 20th century, with agroforestry today taking place only on a few million ha in Europe. (Nair, 2005). In recent years, scientific research, the development of structures and the results of experiments by select professionals has gained recognition of agroforestry and family farming (Lengyel et al., 2016) on national and European levels. Additionally, experimental courses in agroforestry have been launched in several countries, albeit on a small scale, as resources, trainers and available skills are scarce. Fortunately, ICT tools are being continuously developed, so there are tools and methods for e-learning and e-collaboration available for use in the improvement of agroforesty education (Bustos et al., 2007; Herdon and Lengyel, 2013; Herdon and Rózsa, 2012; Szilágyi et al., 2016; Botos, 2012). The main objectives of the development of such tools are to synthesise needs and expectations, based on the current training methods, present training methodologies and present training practices, as well as to set up a common framework to 1) build an innovative training system (contextualized, modularized trainings, use of ICT, with the participation of professionals (Jamnadass et al., 2014)), 2) create a technical collaborative support for the implementation of the project with communication tools (information of partners and promotion) and 3) provide access to the resources and training services during and after the project (knowledge databank, interactive services) (Ibanez et al., 2013). Collaborative Working Environment (CWE) can be perceived as the tools, technologies, services and

[1] Charles Burriel
Agrosup Dijon, France
charles.burriel@educagri.fr
[2] Miklós Herdon
University of Debrecen, Hungary
herdon.miklos@econ.unideb.hu
[3] János Tamás
University of Debrecen, Hungary
tamas@agr.unideb.hu
[4] László Várallyai
University of Debrecen, Hungary
varallyai.laszlo@econ.unideb.hu

environments supporting individual persons in their working tasks to become more creative, innovative and productive (Gunn, 2010) involving the direct or indirect interaction (collaboration) with other individuals, groups or organizations (Collaboration@Work, 2005; Ibanez et al., 2013). Collaborative platforms provide sophisticated upper middleware services, which means that it requires a person-aware distributed collaboration environment (Gregorio et al., 2015). It is based on system integration of Web Services, Semantic Web, CSCW (Computer-Supported Cooperative Work), utility-like computing and connectivity (grid or alike), sensor and wireless technologies (beyond 3G, 4G), advanced network services (e.g. IPv6), knowledge and content management, and WFMS (WorkFlow Management System) based on peer-to-peer design principles to enable radically new collaborative environments. These can provide the support and operations required for complex virtualized working environments. Tasks include development of tools for sharing resources, knowledge/resources discovery, service composition and CSCW tools (including multi-conferencing) to ensure stable, dependable collaborative applications.

2. Objectives

Supporting the agroforestry development (training on different levels (Mbow et al., 2014; Gold et al., 2015), extension activities, farming) in Europe, one important aim of the AgroFE (http://agrofe.eu) and the Agrof-MM (http://agrofmm.eu) European projects is to build a knowledge databank to help supply the stakeholders with agroforestry knowledge. A knowledge base or knowledge bank is a special kind of database (Glick, 2013) for knowledge management. A knowledge base is an information repository that provides the means for information to be collected, organized, shared, searched and utilized. It can be either machine-readable or intended for human use. Behind a Knowledge DataBank (KDB), there is, minimally, a back-end which is based on a special DBMS. Due to the high level of enrichment, a KDB is more than a documentary base or a document management system.

The knowledge management (KM) is used to describe the creation of knowledge repositories, improvement of knowledge access and sharing as well as communication through collaboration, enhancing the knowledge environment and managing knowledge as an asset for the project partners and the public. The central element of the KM is the KDB. In order to build and use the KDB, knowledge engineering is needed. To organize and store the knowledge (represented in different content and forms: e.g., documents, videos, photos), it is necessary to classify them by using knowledge representation technologies. The Dublin-Core (Weibel, 2005; Lubas et al., 2013), LOM (Learning Object metadata), Technical metadata and the Agroforestry thesaurus are used in the project (Castro-García and López-Morteo, 2013; Manouselis et al., 2010).

One crucial part of the development is to define the AgroF-thesauri in every detail, because the words in this vocabulary link to different objects (e.g., document, lecture, presentation, videos) in the KDB. This thesauri system has been developed by a working group in the project using ontology. Unfortunately, the AGROVOC Multilingual agricultural thesaurus and other systems do not contain terms for agroforestry in detail (Rajbhandari and Keizer, 2012).

3. Requirements, functional features

About Knowledge: in the projects, we consider that the “Knowledge” is made of the “digital” parts of the explicit and tacit knowledge related to the Agroforestry domain. The KDB contains the knowledge representation. KDB is the tool of capitalization and ICT are the support of the implementation and uses of the KDB. Knowledge is collective, it does not exist without human beings and the KDB is very partial.

Some important requirements and characteristics of our KDB are the followings:

- Storing elements of knowledge, sources of technical and professional information in relation to the domain;
- Making professional experiences accessible to learners, students, farmers in training, experiences under the form of document, reports, video, evidence and keynotes;

- Allowing the diffusion of knowledge in the form of apprenticeship modules, together or linked to education and training platform;
- To be one of the support tools of the collaborative and interactive work in the project, especially as a means to help the production of professional targeted knowledge;
- Bringing agroforestry knowledge content and bridging the other related KDBs (agronomy, agro ecology, environment, water, organic farming) and agroforestry sources;
- In mid-term, for the "public part" of the KDB, making this knowledge accessible to the public, in respect to intellectual property.

4. Results

4.1. The developed system

The knowledge database has been developed to be used as tools and training resources and will also integrate with existing and future training resources. Collaborative and dissemination platforms were created such as an official web site (http://agrofmm.eu/index.php/en/), twitter (https://twitter.com/AgrofMM) a video-conference system and Facebook page have been registered and maintained (https://www.facebook.com/AgrofMMEU/), a mailing list and a Moodle training portal for project documentation (http://moodle.agrofmm.eu/) and as a Learning Management System (LMS) (Detterbeck and Sciangule, 2017). In the context of the AgroFE project, the technologies of information and communication, (ICT) include four components: 1) the collaborative tools (CMS – Collaborative Management System); 2) the Knowledge DataBank (KDB) based on RUBEDO (a developer system https://www.rubedo-project.org/en/), the MONGODB (a non-relational database handling system) and an Elastics Search engine, 3) Editorial chain and 4) E-learning platform for training and archiving. The KDB is to enable the sharing, access and consultation in the use of certain resources for training (Figure 1).

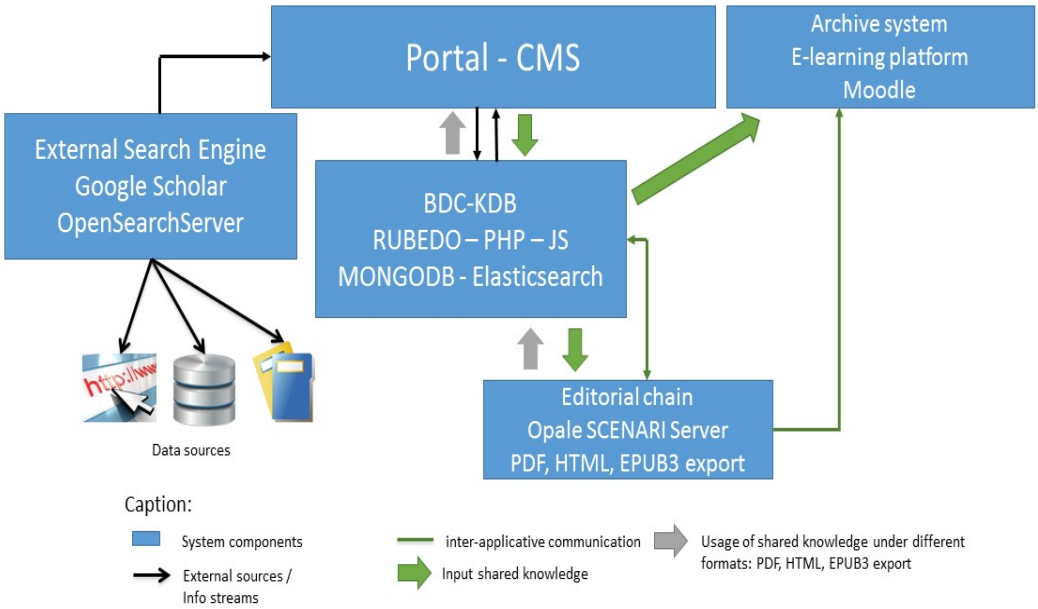

Figure 1. The integrated service and knowledge base system architecture (The concept created by Charles Burriel)

The knowledge databank is a component of the project training system. It aims to gather and share a set of documents, resources that partners can use and which learners and public users can access. The project focuses on the newest innovative ICT solutions and trends. The knowledge databank enables access to, sharing of and consultation on the resources for training. These resources are in different forms: Mono document object (like a photo, a text, a diagram) and Composite materials (for example a html web page with images, a pdf file with pictures and diagrams, a video clip with images and sounds ...).

The knowledge databank is based on the RUBEDO system developed in PHP. RUBEDO is built on different components: a database management software (DBMS), NoSQL, MONGODB (Abbes and Gargouri, 2016); the user interface uses the Elasticsearch search engine (https://www.rubedo-project.org/en/). The RUBEDO system transforms the digital experience of users, bringing answers to the current web challenges:

- real time personalization;
- multi-sites;
- mobility;
- performance.

The RUBEDO platform (Webtales, 2017) brings the advanced functions of content management together with the power of e-commerce, all being based on Big Data.

The system includes tools to design the appropriate pages for the users. These pages can be seen by users on the FrontOffice part in the different kind of browser program. The visitors, learners and trainers can access the KDB content on the http://kdb.agrofmm.eu URL without further specification. The language of the user interface can be changed by clicking on the language button and by selecting the appropriate language, (i.e., French, English, Hungarian) from the pull-down menu. The page is redisplayed in the selected language. Details about access, consultation and research are provided in the pages of Welcome and Oveview menus. The general menu which contains Trainings, Resources, Content types and Search tool are on every page. The KDB contains different media type files which the Digital Assets Manager module of the RUBEDO.

The objective of the KDB is to collect, store and share certain resources for training helping the trainers, advisors, learners and farmers in their tasks. These resources are in different forms:

- Mono document object, like a photo, a text or a diagram.
- Composite materials, for example an html web page with images, “pdf” files with pictures and diagrams, video clips, sounds files, etc.

During the project, these documents are identified, selected, proposed by partners and included into the KDB for the evaluation of their potential use in training, by one or more partners. A fact sheet originally written by the proposer often accompanies and completes this document.

The KDB is based on different professional vocabulary, metadata and thesaurus system that are used for building the content structure and helping the users in searching. A new tool will improve knowledge management of agroforestry. The Agrof-MM project is creating an agroforestry thesaurus. This thesaurus is a glossary which will include all terminology related to agroforestry, such as agroforesters, agroecology, soil protection, as well as the relationships between these terms and others. Currently under development by scientists, teachers and agroforestry experts, the glossary will be in the language of each partner country.

4.2. Access: Taxonomy and Enrichment

For designers and developers, the design and developments associated with the management and use of the KDB covers four dimensions: the “professional domain” dimension composed of Vocabulary, Thesaurus, Taxonomy of agroforestry; the “training” dimension here dealt with the question of the enrichment of documents; the “usage” dimension, with its 2 components: access based on enrichment and the user interface, the “users” dimension, which links the first 3 dimensions, mentioned before.

Easier access and pathway into KDB content: the object-documents are enrichment. In relation to KDB access, different modes and levels of enrichment have been or will be developed and implemented. The search tool can be accessed as Front Office interface, one example of the searching resultcan be seen on Figure 2. The system can be used by searhing with taxonomy based terms and/or by using expression with full text search. On the Figure 2 on the search field the “agroforestry and burriel” search experession and on the left hand side different taxonomy attribute can be seen. The users can select one or more of them and the taxomomy system gives possibilities to offer more

thesaurus system. For example Dublin Core, Learning Object Metadata and Agroforestry terms will appear for selection.

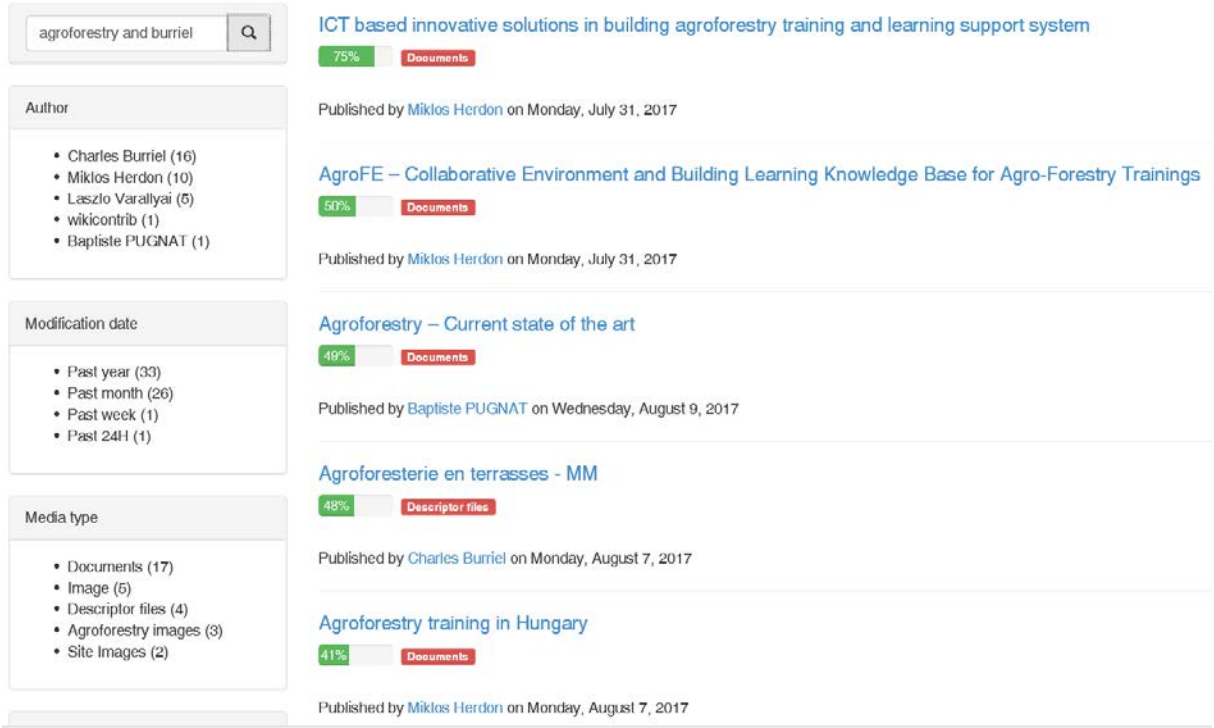

Figure 2. The Elasticsearch in the Front Office programme (Screenshot of KDB search tool)

Four access levels, A1, A2, ... have been identified and different enrichment modalities E1, E2, ... were successively developed.

1) A1: internal identification of the object, the document representing the knowledge, starts with the file name plus internal identifier (key, UID…), admin management under the system associated with the DBMS Mongo dB; A1 => E1: only for access-management by the administrators, only "system" enrichment.

2) A2: "simple" access-consultation, by keywords resulting from the indexing by the Elasticsearch full text engine of the RUBEDO platform; A2 => E2: "technical" enrichment by indexing engine, incomplete taking into account the types of objects (Photo, documents, video, …) that can be consulted, consequently, considered as insufficient.

3) A3: First level of human enrichment (E3a and E3b) for access to documents / objects: additions of keywords. This level of enrichment allows us to complete the information (metadata) of objects like photos, graphics.

   Two evaluated solutions:

   - E3a: Added metadata and synonyms in the document, like metadata of a scientific article that lists keywords at the end of the summary on page 1 of the article;
   - E3b: Association of the document entered into the KDB with 4 taxonomies (Domain, DC – Dublin Core, LOM – Learning Object Metadata, TECH – Technical Metadata) usually allows sophisticated query development. However, the Domain taxonomy was "primary", simplified in the v2 because there is no Vocabulary-Thesaurus-Taxonomy (VTT) for Agroforestry, either in France or at the European level (Figure 3). Figure 4 and Figure 5 shows the taxonomy management in the Back Office of the KDB system.

4) A4: As part of the development of the KDB v3, the A3 process is redesigned, simplified / re-developed and a new educational enrichment will be implemented. It is in the conception phase

at the time this abstract is being written. The A3 level of v3 integrates the development of a VTT based on the work of specialists.

- E4a takes over the E3a process, with comparative assessment of on-line direct typing and offline input;
- E4b will rely on only 2 taxonomies, VTT and LOM.
- E4c, Educational enrichment; It will include a description of the document's components (URLs linking components), descriptors of the document's component(s), compatible with taxonomy Domain and LOM.

Consequently, the development of a VTT is part of the work of KDB version 3.

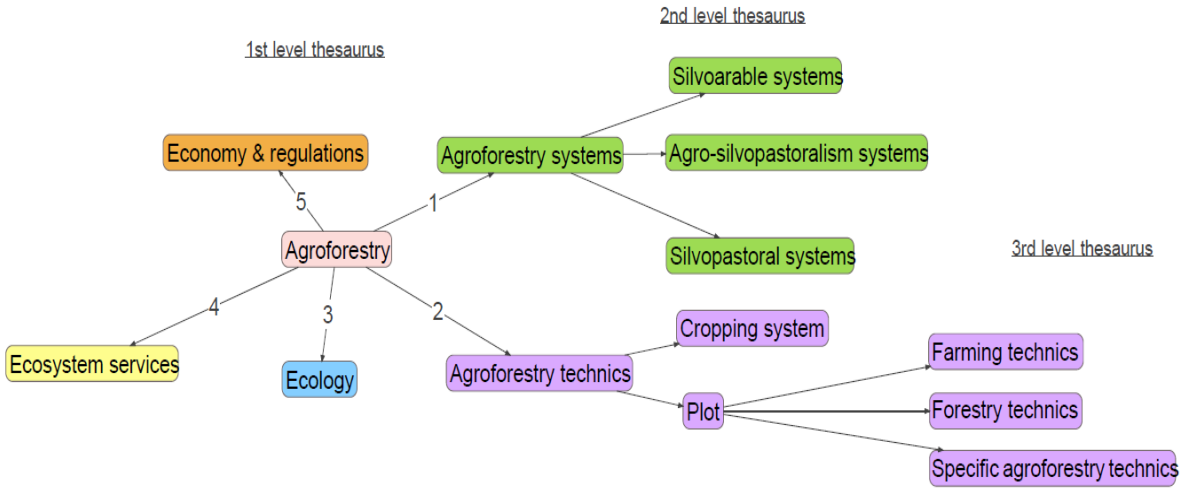

Figure 3. The levels of the thesaurus system for the Agrof-MM KDB (Develeoped by one of the Agrof-MM working group and composed by Jean-Michel Escurat)

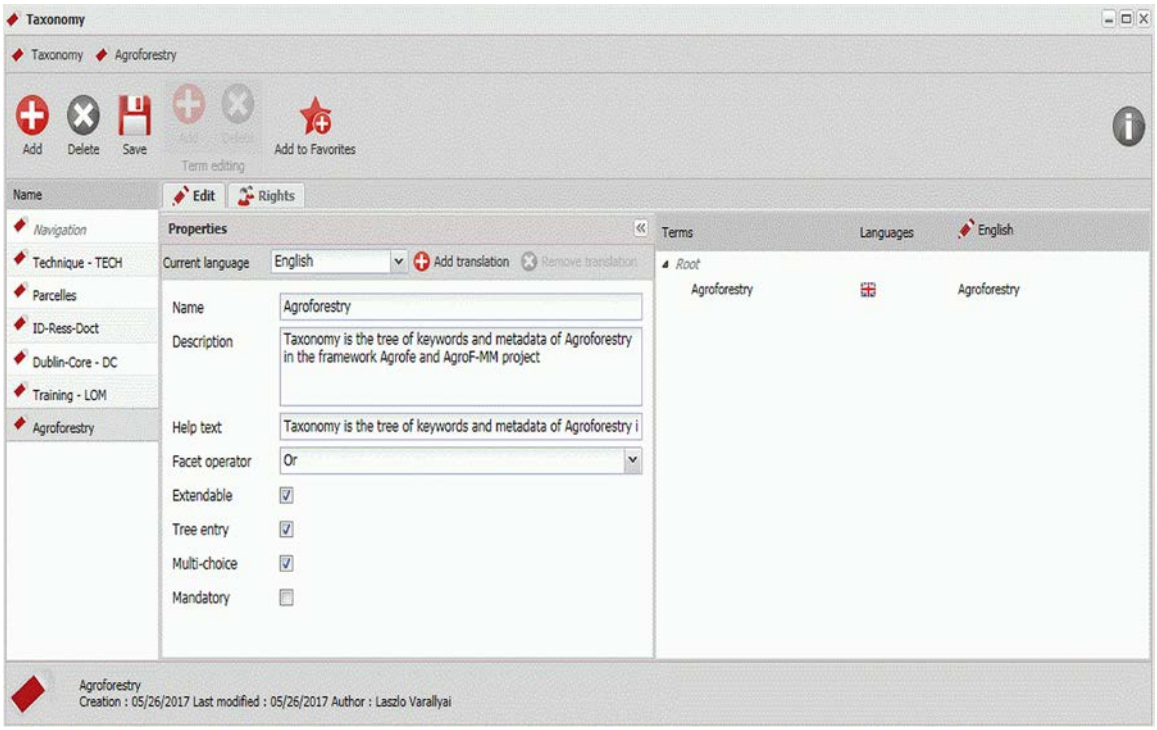

Figure 4. The Taxonomy definition in Rubedo system (Taxonomy management function in the KDB - Screenshot)

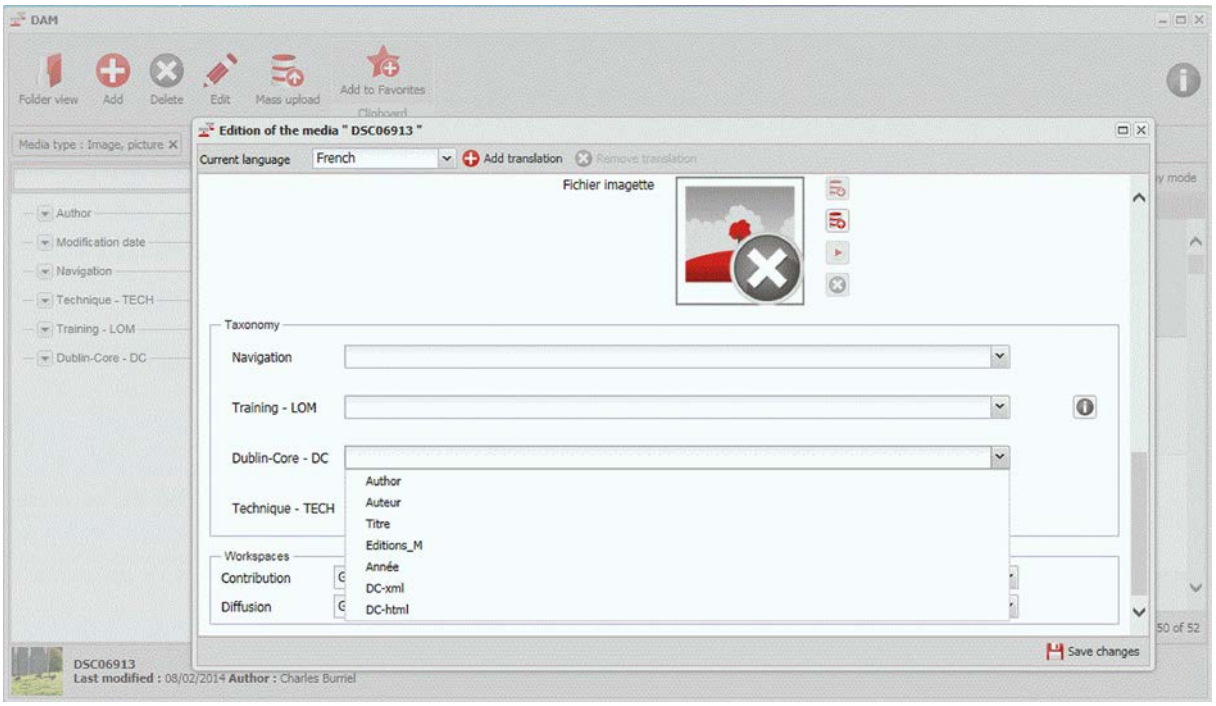

Figure 5. Assign the taxonomies (LOM, DC, TECH) to the image in the Rubedo system (KDB Back Office screenshot)

5. Evaluation - validation - conclusion and perspectives

The validation process of the KDB v3 will involve the participation of all potential users. The opinion of computer scientists and ergonomists will also be enquired as well as opinion from the public. These evaluations will guide the evolution of the KDB, the extension of uses the knowledge database, the integration of the documentary base(s) for professional purposes or the extension to other fields associated with agroforestry, biodiversity for example.

The knowledge database (knowledge data bank) is a new innovative solution for harvesting, storing and delivering contents in agroforestry. It was used in different training programs with positive feedback. The knowledge database will serve the Agrof-MM partners in the next years.

Acknowledgements

This publication was supported by EU Leonardo Innovations Transfer 'Agroforesterie Formation en Europe – AgroFE' Ref. Number: 2013-1-FR1-LEO05-48937 project and EU Erasmus + Programme Key Action 2: Strategic Partnership. 'Agroforestry – Training – Mediterranean and Mountain' Ref. Number: 2015-1-FR01-KA202-015181 project.

References

Abbes H., Gargouri F. 2016. Big Data Integration: A MongoDB Database and Modular Ontologies based Approach, Procedia Computer Science, Volume 96, p. 446-455. https://doi.org/10.1016/j.procs.2016.08.099

Botos Sz. 2012. Broadband development: a case study in a rural micro-region. SEFBIS Journal 2012 (7) p. 56-65.

Bustos E.S., F.S. Zazueta, H. B. Howard. 2007. Rapid prototyping of learning objects and their implemetation using ontology editor. In Conference on Agricultural Economics, Rural Development and Informatics, pp. 59-60. Debrecen, Hungary.

Castro-García L., López-Morteo G. 2013 An international analysis of the extensions to the IEEE LOMv1.0 metadata standard, Computer Standards & Interfaces, 35(6) p. 567-581. https://doi.org/10.1016/j.csi.2013.04.006

Collaboration@Work. 2005. The 2005 Report on new working environments and practices. Luxembourg: Office for Official Publications of the European Communities, 2005 ISBN 92-894-9283-X. 2005.

Detterbeck K., Sciangula M. 2017. Chapter 12 - Innovation Through Collaboration: Using an Open-Source Learning Management System to Enhance Library Instruction and Student Learning, Distributed Learning, pp. 221-238. https://doi.org/10.1016/B978-0-08-100598-9.00012-X

Glick J. 2013. Chapter 8 - Ontologies and Databases – Knowledge Engineering for Materials Informatics. Informatics for Materials Science and Engineering, pp 147-187. https://doi.org/10.1016/B978-0-12-394399-6.00008-4

Gold M., Cernusca M. and Hall M. (Eds.) 2015. Training Manual for Applied Agroforestry Practices. The Center for Agroforestry. University of Missouri. http://www.centerforagroforestry.org/pubs/training/

Gregorio, N., Herbohn, J., Harrison, S. and Smith C. 2015. A systems approach to improving the quality of tree seedlings for agroforestry, tree farming and reforestation in the Philippines. Land Use Policy, (47), p. 29-41. https://doi.org/10.1016/j.landusepol.2015.03.009

Gunn C. (2010) Sustainability factors for e-learning initiatives. Research in Learning Technology, 18 (2), p. 89–103. https://doi.org/10.1080/09687769.2010.492848

Herdon M. and P. Lengyel. 2013. Building and Using Knowledge Repositories for Agriculture: An Innovation Case Study. In: EFITA WCCA CIGR 2013 Conference: Sustainable Agriculture through ICT innovation, pp. 180-187. Torino, Italy.

Herdon M. and T. Rózsa. 2012. Knowledge dissemination on innovative information technologies in agriculture. Journal of Ecoagritourism 8:(1), pp. 301-306.

Ibanez M.B., Rueda J.G., Maroto D. and Kloos C.D. 2013. Collaborative learning in multi-user virtual environments. Journal of Network and Computer Applications, 36:(6), pp. 1566-1576. https://doi.org/10.1016/j.jnca.2012.12.027

Jamnadass, R., Langford, K., Anjarwalla, P. and Mithöfer D. 2014. Public–Private Partnerships in Agroforestry. International Journal of Educational Development 16(1), p. 544-564. https://doi.org/10.1016/B978-0-444-52512-3.00026-7

Lengyel P., Pancsira J., Balogh P., Oláh J., Füzesi I. 2017. Social network analysis of international scientific collaboration on family farming research. Journal of Agricultural Informatics 8:(2) pp. 71-79.

Lubas R.L.; Jackson A.S.; Schneider I. 2013. Using Dublin Core: With information from the Dublin Core Metadata Initiative, The Metadata Manual A Practical Workbook, p.45-65. https://doi.org/10.1016/B978-1-84334-729-3.50003-0

Manouselis N., Najjar J., Kastrantas K., Salokhe G., Stracke C.M.,Duval E. 2010. Metadata interoperability in agricultural learning repositories: An analysis, Computers and Electronics in Agriculture, 70(2) p. 302-320. https://doi.org/10.1016/j.compag.2009.07.007

Mbow, C., van Noordwijk, M., Prabhu, R. and Simons T. 2014. Knowledge gaps and research needs concerning agroforestry's contribution to Sustainable Development Goals in Africa. Current Opinion in Environmental Sustainability. (6), p. 162-170. https://doi.org/10.1016/j.cosust.2013.11.030

Nair, P.K.R. 2005 Agroforestry. Encyclopedia of Soils in the Environment. p. 35-44.

Rajbhandari S., Keizer J. 2012. The AGROVOC Concept Scheme – A Walkthrough, Journal of Integrative Agriculture, Volume 11(5) p. 694-699. https://doi.org/10.1016/S2095-3119(12)60058-6

Szilágyi R., Lengyel P., Füzesi I. 2016. ICT in Agriculture in Aspect of Regional Differences. Journal of Ecoagritourism 32:(1) pp. 30-34.

Webtales. 2017. http://www.rubedo-project.org/en/homepage/rubedo-dream-team (15.05.2017).

Weibel S.L. 2005. The Dublin Core Metadata Element Set ISSN: 1041-5653, World Spatial Metadata Standards, p. 493-513.