

REPOSITORY CONVERSATIONS

Ana Alice Baptista
Miguel Ferreira
University of Minho

ABSTRACT

For a long time, informal communication has played an important role in the process of creating and sharing scholarly knowledge. However, only a handful of initiatives have seriously attempted to exploit the advantages of informal communication processes as a complement to the classic mechanisms of scholarly communication. This article describes the developments that resulted from the DSpace-dev@University of Minho project, particularly, the ones that focus the Commenting add-on: an extension developed for the DSpace platform which aims at providing Institutional Repositories with informal communication capacities. The article also alludes to some ideas that can be taken as the basis for future developments about this subject, namely: to transform the current add-on into a cross-repository and cross-platform service; to make each comment have its own descriptive metadata in order to facilitate its discovery by search engines; and to improve significantly its graphical user interface.

Key-Words: Institutional Repositories; Scholarly Communication; Informal Communication; DSpace.

INTRODUCTION AND FUNDAMENTALS

"It is worth remembering [...] that communication changes have always occurred: it is simply that current changes are particularly rapid and radical".
Meadows, 1997

Scholarly communication 'is the system through which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community and preserved for future use. The system includes both formal means of communication, such as publication in peer-reviewed journals, and informal channels, such as electronic listservs' (ACRL, 2003).

The importance of informal communication in the overall scholarly communication process is frequently neglected. Nonetheless, studies have firmly demonstrated that its role is highly preponderant both during the creation phase of new

scholarly knowledge, as afterwards, being a prime facilitator for its dissemination. (EHRlich; CASH, 1994; LEVY; MARSHALL, 1994; MARSHALL, 1997, 1996; MARSHALL et al., 1994).

The so-called 'social software' creates a whole new framework around these questions. Tools such as del.icio.us¹, Connotea² or Flickr³ enable everyone to carry out an assortment of activities such as adding informal comments to items of information, indexing those items using free terms⁴, describing Web content in a collaborative way or in a stand-alone environment, etc. However, one must note that even though some of these tools allow the user to 'attach' a comment to an item of information, is not always true that this will be done as a support for discussion. It is not possible, for example, to generate discussions or threads based on an existing document on the Web. They are often used as ways to express an opinion about a certain Web resource, and it is not possible for someone to re-comment a comment generating a conversation thread about a certain subject.

On the other hand, commenting functionality provided by this type of software only enables the user to attach a comment to a sole Web resource. Nevertheless, when people speak freely, far from technological constraints, they often relate information coming from various different sources, i.e. they compare articles, methods and approaches, they complement information found in an article with information from other articles, etc.

It is also important that comments and other types of informal communication are more 'visible' to search engines and, subsequently, to end-users. Some of the strategies to achieve that goal can be acquired through a direct relationship with the formal documents and/or with the inclusion of metadata elements within the comment it-self.

The usual informal communication tools (e.g. electronic mail) are being used in the context of scholarly communication right since beginning of its existence. The

1 See <http://del.icio.us>

2 See <http://www.connotea.org>

3 See <http://www.flickr.com/>

4 To this form of freely and collaboratively indexing the contents of the Web it is given the current name of *folksonomy*.

Web itself appeared in a scholarly communication environment. Social software is also being used in this context. Connotea is one of the tools in this category more strongly directed at being used by scientists (HAMMOND et al., 2005).

Since most of these tools are general enough to be applicable in many different contexts, it is important to identify the ones that are more suitable to be used in scholarly communication contexts. This specific type of communication has particularities that are hardly found in other communities. For example, in these environments people often perform, simultaneously, the roles author and reader.

At the same time, the scholarly communication process is undergoing deep changes. The Open Access (OA) movement is responsible by many of biggest current changes. OA means “free, immediate, permanent online access to the full text of research articles for anyone, worldwide” . There are two ‘roads’ for OA: the ‘green road’ and the ‘golden road’. The ‘green road’ is “where authors provide OA to their own published articles, by making their own eprints free for all”; the ‘golden road’ is “where journals provide OA to their articles” (HARNAD).

An University-based Institutional Repository (IR) ‘a set of services that an university offers to the members of its community for management and dissemination of digital materials created by the institution and by its community members’ (LYNCH, 2003). The IRs are generally used to implement the ‘green road’ of OA.

Although the proposals presented in this article may be generally applied to all sorts of scholarly communication, here they specifically refer and are instantiated in the context of IRs, as it is envisioned that in the near future nearly all scholarly literature will be available in OA through IRs.

During the Dspace-dev@University of Minho⁵ project, four add-ons were developed for the Dspace platform: the controlled vocabulary add-on⁶ (FERREIRA; BAPTISTA, 2005), the commenting add-on⁷, the recommendation add-on⁸, and web of

⁵ See <http://dspace-dev.dsi.uminho.pt>

⁶ See http://dspace-dev.dsi.uminho.pt/pt/addon_ontology.jsp

⁷ See http://dspace-dev.dsi.uminho.pt/pt/addon_commenting.jsp

⁸ See http://dspace-dev.dsi.uminho.pt/pt/addon_recommendation.jsp

communication add-on⁹. In the following sections, we will describe in detail the commenting add-on and highlight the changes it needs in order to leverage the importance of scholarly informal communication.

COMMENTING ADD-ON

The commenting add-on allows users to attach textual comments to DSpace resources. Although it appears as a regular newsgroup application, this add-on implements a graph-based approach to comment threads. This is a very distinctive aspect of this add-on because it makes a huge difference in the way resources are related. While in a tree-based organization, postings can only be attached to a single resource, in a graph-based approach comments can simultaneously be attached to several different resources (Figure 1).

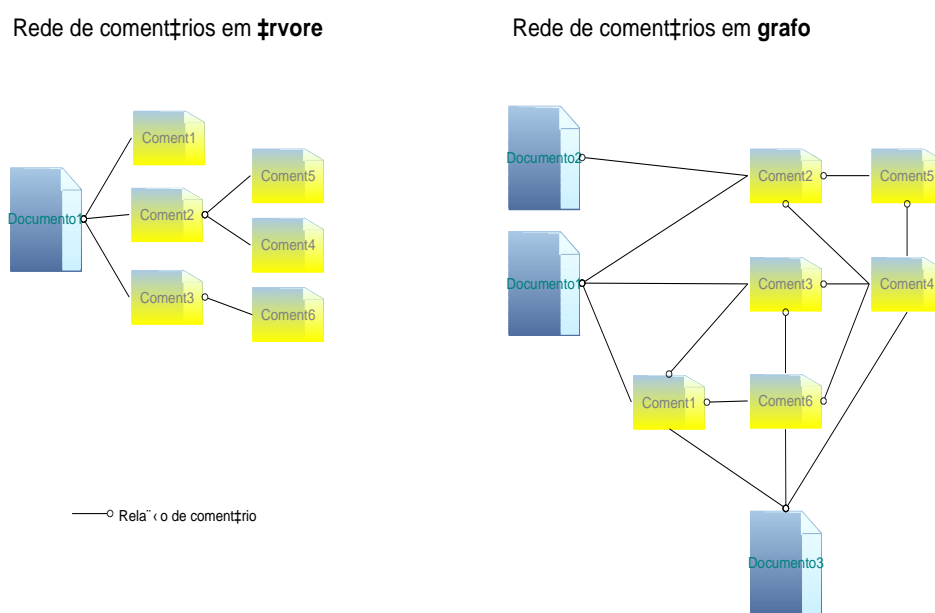


Figure 1 - Difference between the commenting net in trees and in graphs.

The theory that served as a basis for the construction of this add-on is that each formal document (a journal paper, for instance) implies the creation of a knowledge

⁹ See http://dspace-dev.dsi.uminho.pt/pt/addon_webcommunication.jsp

community composed by the people that read that very same document. Some of these people may refer this document in other formal documents they produce, but the majority will just talk and share ideas about it at the informal level. In these conversations, true pearls of wisdom may be exchanged, that if not properly stored for future use, will only remain with the people who were directly involved in that communication process (BAPTISTA et al., 1999).

In this approach, basically there are two different levels of scholarly communication: (1) **the formal communication level** that consists of documents that are published by scholars, and (2) **the informal communication level** that consists of informal documents that, in this case, are the comments both to formal and informal documents. The informal communication level is generated from the formal communication level, and it is inwardly connected to it (Figure 2).

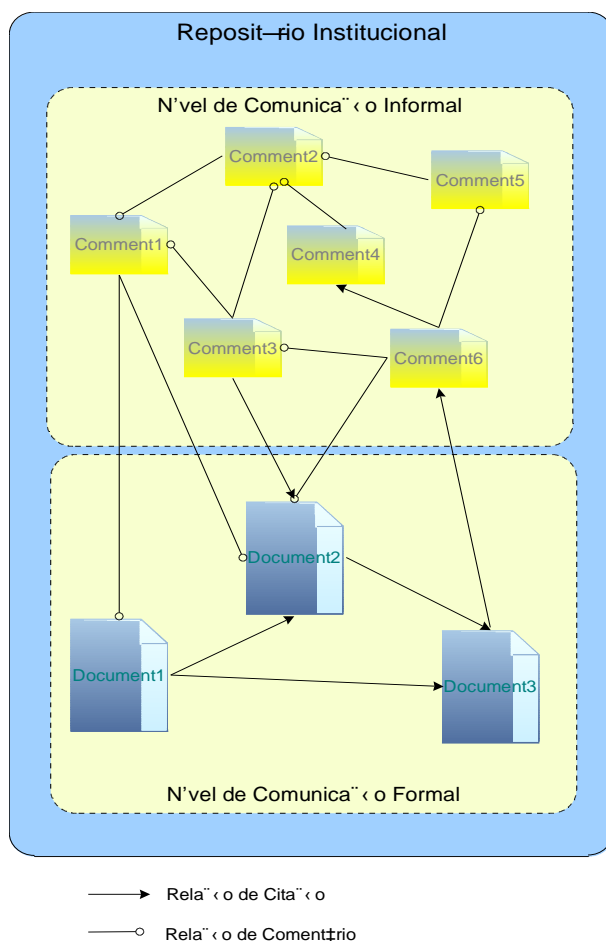


Figure 2 – Relationship between the formal and informal communication levels provided by

the commenting add-on to DSpace.

CHANGES TO BE MADE TO THE COMMENTING ADD-ON

The current version of the commenting add-on has some limitations that we intend to overcome:

1. **It is local to a repository** – It is only possible to comment resources within a specific instance of DSpace (i.e. the one where the add-on is installed). The way formal documents are related is independent of the repositories where they sit (i.e. it uses the classic citation system). On the other hand, under an informal perspective scholars relate information stored in many places – they are not limited to the place where they have their documents stored or the place where they work. Thus, the commenting functionality should be cross-repository, as it happens to the majority of available social software tools;
2. **It only works with the Dspace platform** - The add-on was developed specifically for DSpace. However, it would be more interesting to have an informal communication service that could be used independently of the platform where the formal resources are stored. Although such a service could be applied to the Web as a whole, it is important to consider the specificities of the scholarly communication and emphasize its necessity in this environment where the readers are also the authors. In the current context of transition for a 100% OA environment, it makes sense that such a system should be applied at least to the documents stored in IRs;
3. **The comments are not indexed under any subject** – It is important that users are able to include descriptive terms coming from controlled vocabularies and/or free keywords in their comments. It is necessary to follow attentively the present trends in free tagging that lead to self-

organized communities, such as the ones that emerge naturally in social software environments (folksonomies);

4. **Limited graphic user interface** – The current design of the graphic user interface of the commenting add-on is poor both in the selection of documents to be commented and in the presentation of the network of comments. Figure 3 shows a screenshot of a comment thread. Observe that the comments 3 and 6 appear twice. This happens because the presentation is made in the traditional way, as if the commenting net was organized in trees, not graphs. Changes to this add-on should include the contribution from graphical design specialists for the Web.



Figure 3 - Screenshot of the graphic feature of a comment thread.

The resulting service will have an architecture similar to the one on Figure

4.

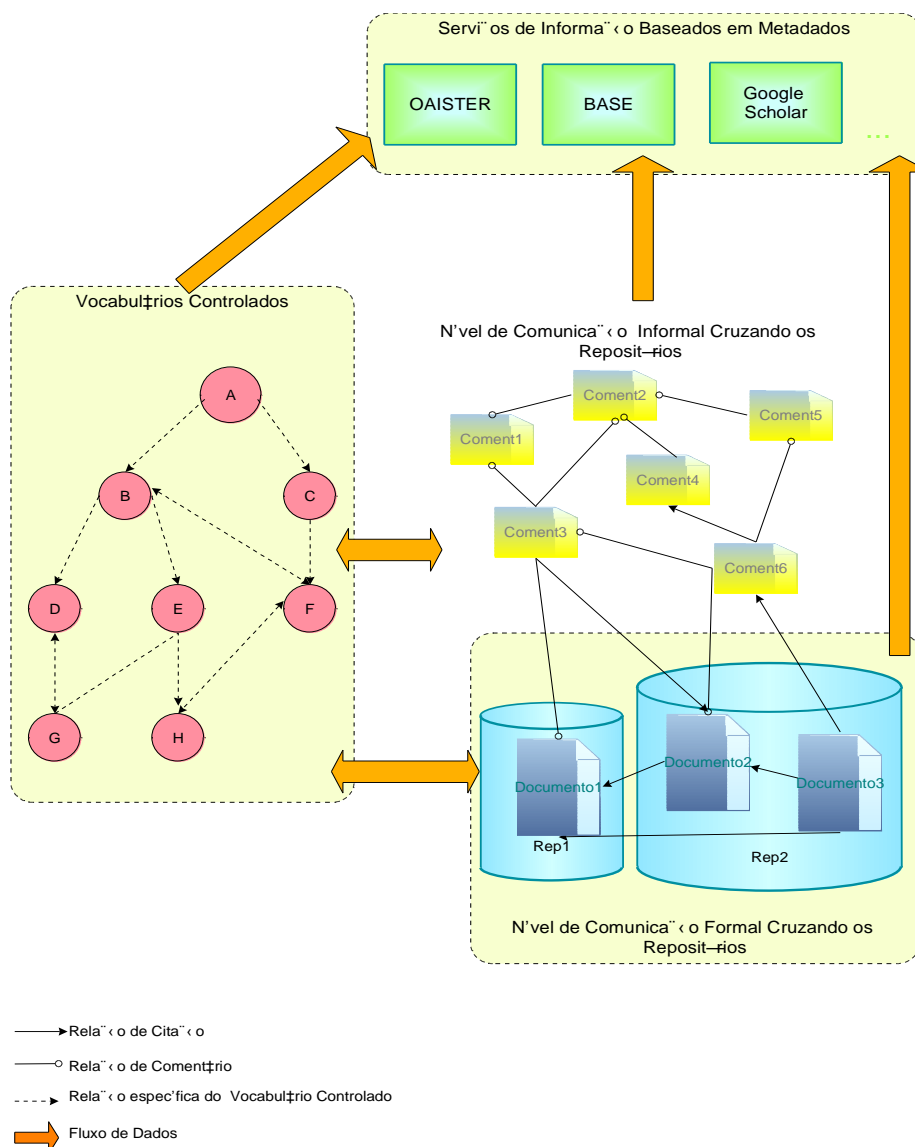


Figure 4 - Architecture of the new comment service.

After two years of experience with the commenting add-on being used on Papadocs¹⁰, it seems that the users have some difficulty in commenting the works of others. We do not have any evidence to confirm this assumption other than the fact that there are not many comments being posted on the platform. This lead us to believe that the users are not comfortable with the graphical interface provided by the add-on, or they have a feeling of awkwardness when commenting the results of the work of their peers. If

¹⁰ See <http://papadocs.dsi.uminho.pt>. Papadocs is not an IR – it is a DSpace instance, but it is used to store students' assignments so that they can be available for all their peers and teachers of the following years. It can be also used to generate the individual portfolio of each student.

so, then some sort of intermediate service could aid solving this problem.

A service that could act at the intermediate level could be a personalized blog. Blogs have the advantage of having acquired the social habit of commenting by its users. Because blog posts are, most of the times, considered as informal communication and comments are organized sequentially and unrelated to each other, some changes would have to be introduced:

1. **Posts** – Blog posts would be generated directly from article abstracts, i.e. each abstract would automatically generate a post in the blog;
2. **Access** – The blog could be personalized: each user would virtually access a different blog according to his/her profile. This blog would be, in fact, a view of the documents stored in the global net of repositories;
3. **Comments** – The comments would be made in the blog environment, but organized as a graph. The commenting network generated from the post would be attached to the formal document that originated it. The user would be able to chose the right view for him/her: a document view or a blog view. The principle would be exactly the same: one or more formal documents related through informal documents. The rest could be totally artificial and generated only to provide better user's adapted environment.

Figure 5 shows the architecture for the new comment system.

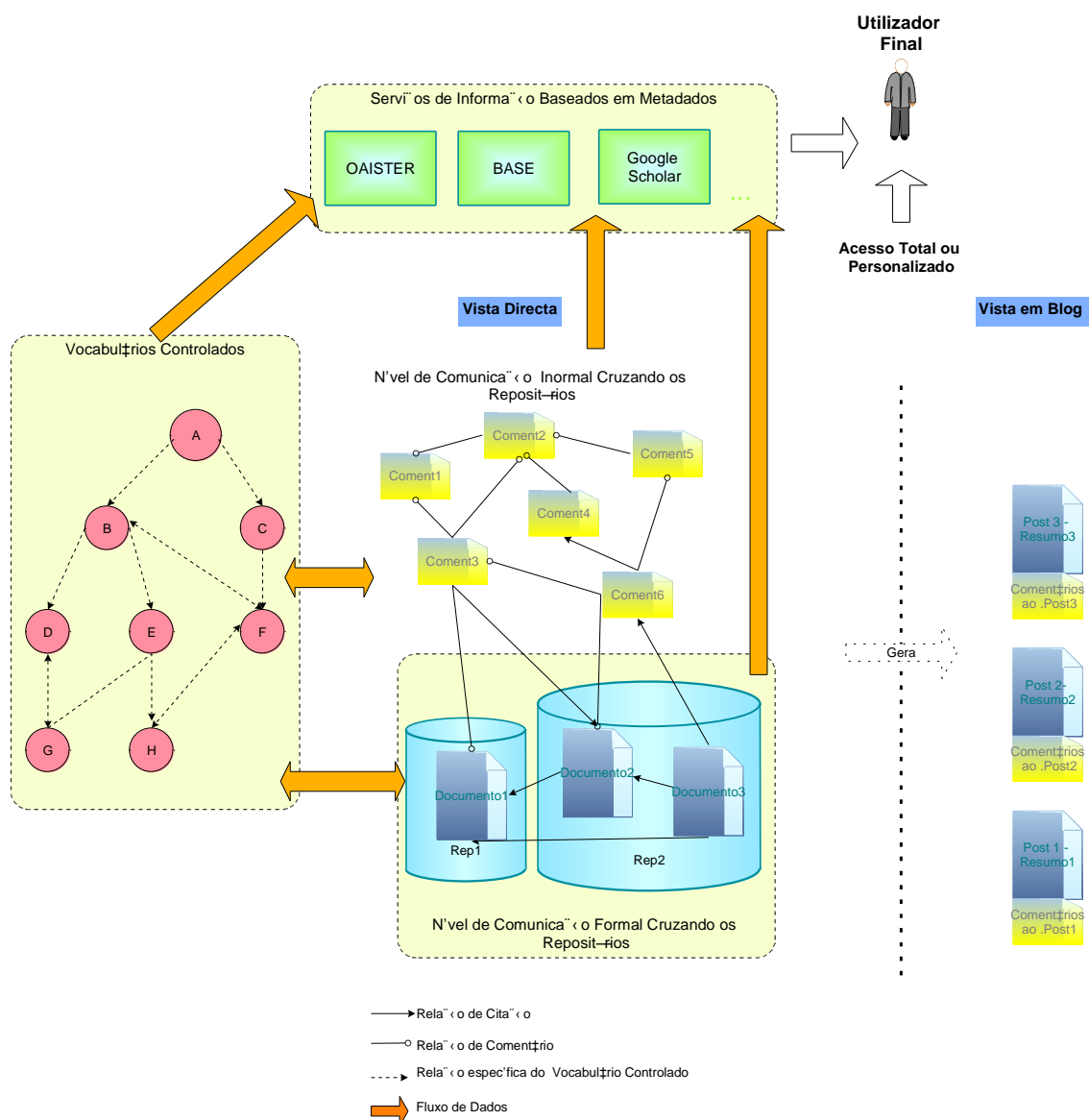


Figure 5 – Final architecture of the new comment system

As referred in the introduction three other add-ons were developed under the scope of the DSpace-dev@University of Minho project. The Web of Communication (WoC) is one of them.

A Web-of-Communication (WoC) is graphical representation of the network that emerges from the exchange of items of information among people, i.e. communication. In scholarly communication, these people are scientists and academics that, as previously mentioned, act simultaneously as authors and readers of scholarly material. The construction of this add-on was based on the assumption that it is possible to

interpret and graphically represent the communication relationships between people if one can have access to the records of the communication (both formal and informal) interactions that occur between them.

Within institutional repositories there is a large amount of formal records. When informal communication capabilities are added to the repository and some requirements of the metadata (e.g. author's name, subject, keywords, description) are assured, a large amount of rich informal communication records is generated. All this information can be used as an input for the WoC.

Figure 6 shows a screenshot of the current WoC add-on graphical user interface.

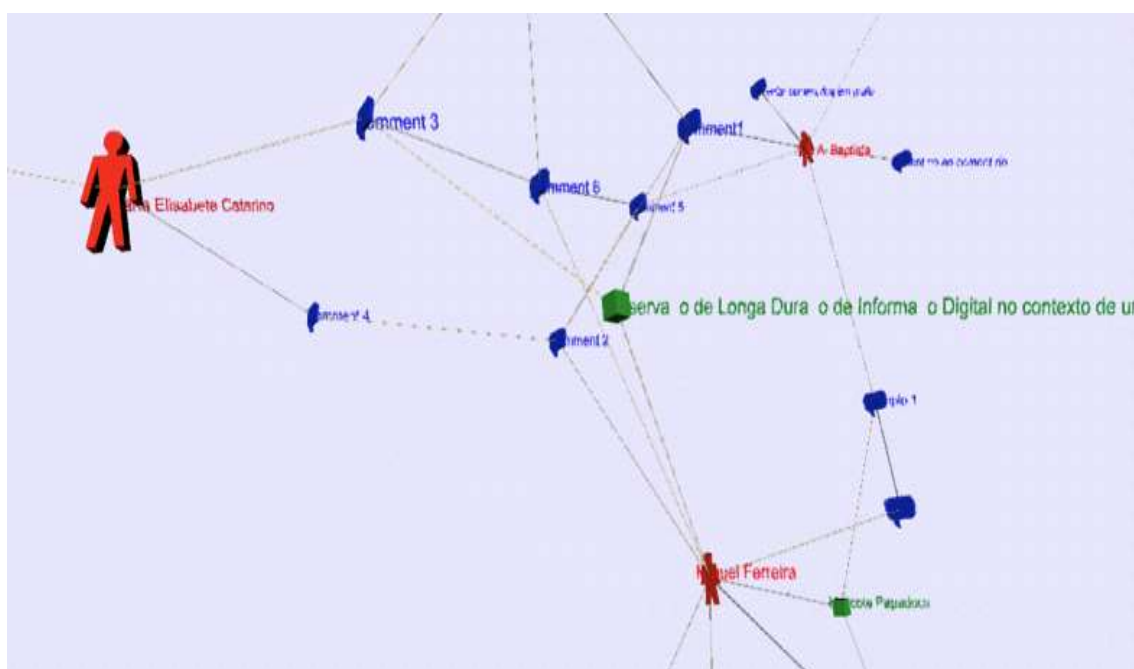


Figure 6 - Print screen of the present add-on in the web of communication

This add-on will also be changed in order to map communication relationships between the scientists and academics in a more reliable way¹¹.

¹¹ An article will soon be published containing further details on this subject.

CONCLUSIONS AND FUTURE WORK

Four add-ons for DSpace were developed in the context of the DSpace-dev@University of Minho project: the Controlled Vocabulary add-on, the Commenting add-on, the Recommendation add-on, and the Web-of-Communication add-on.

The Commenting add-on is particularly important as it is responsible for adding an informal communication layer to DSpace. Basically, this add-on allows comments to be added to DSpace resources. These comments are organized as graphs as it is possible to simultaneously comment several DSpace resources (i.e. other comments and/or formal documents) with a single comment. .

Considering the experience acquired over the last years and the state-of-the-art in terms of Web tools, we believe it is now time to introduce some changes to this add-on. First of all, it will be important that the add-on is transformed into an independent system that should be cross-repository and cross-platform. It will be also of great importance to have metadata regarding comments that could then be used in Semantic Web contexts, e.g. subject, type, or the author's name. This metadata will also be of great importance to improve the WoC add-on, as it tries to map the communication relationships that exist between people. In this context, at least the author's name, document's subject, and document's type may be of fundamental importance.

In order to promote the commenting activity in institutional repositories, we propose the creation of a personalized and customizable blog. Posts in that blog will be article abstracts. Comments attached to blog posts will be, in fact, considered as comments to the formal documents that originated the posts.

An international team is currently being created in order to develop a service that combines the features that support all the add-ons developed in the scope of the DSpace-dev@University of Minho project, and extends them to be used beyond the domain of a local repository.

Some of the new functionalities of this new system are the ones described in this article. Other functionalities are already being determined, and are still to be published. Others will result from studies that will be accomplished after the work team is formally constituted.

REFERENCES

ACRL, S. C. C. (2003, 24 June 2003). **Principles and strategies for the reform of scholarly communication**. Retrieved: 18 October, 2006. Available at: <<http://www.ala.org/ala/acrl/acrlpubs/whitepapers/principlesstrategies.htm>>.

BAPTISTA, A. A., et al. (1999). **Online publishing as a support for scholarly communication in dynamic knowledge communities**. Paper presented at the Electronic Publishing '99: redefining the information chain: new ways and voices., Ronneby, Sweden. Available at: <<http://hdl.handle.net/1822/169>>.

EHRlich, K.; CASH, D. (1994, June). **Turning information into knowledge: information finding as a collaborative activity**. Paper presented at the The First Annual Conference on the Theory and Practice of Digital Libraries, Texas - USA. Available at: <<http://csdl.tamu.edu/DL94/paper/lotus.html>>.

FERREIRA, M.; BAPTISTA, A. A. (2005). **The use of taxonomies as a way to achieve interoperability and improved resource discovery in DSpace-based repositories**. Paper presented at the XATA 2005 : XML : aplicações e tecnologias associadas, Braga. Available at: <<http://hdl.handle.net/1822/873>>.

HAMMOND, T., et al. (2005, April, 2005). Social bookmarking tools (I). **D-Lib Magazine**, 11. Available at: <<http://www.dlib.org/dlib/april05/hammond/04hammond.html>>.

HARNAD, S. **Open Access**. Retrieved: 14/12/2006, 2006. Available at: <<http://www.eprints.org/openaccess/>>.

LEVY, D. M.; MARSHALL, C. C. (1994, June). **Washington's white horse? A look at assumptions underlying digital libraries**. Paper presented at the The First Annual Conference on the Theory and Practice of Digital Libraries, Texas - USA. Available at <<http://csdl.tamu.edu/DL94/paper/levy.html>>.

LYNCH, C. A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. **ARL Bimonthly Report** 226. Available at: <<http://www.arl.org/newsltr/226/ir.html>>.

MARSHALL, C. C. (1997, July 23-26, 1997). **Annotation**: from paper books to the digital library. Paper presented at the Second ACM International Conference on Digital Libraries, Philadelphia, Pennsylvania - USA. Available at: <<http://www.csdl.tamu.edu/~marshall/dl97.pdf>>.

MARSHALL, C. C. (1996). **Awareness technologies and the corporate Intranet**: compelling arguments for adoption, convincing tales of tribulation. Paper presented at the CSCW' 96. Available at: <<http://www.csdl.tamu.edu/~marshall/cscw96-workshop-paper.html>>.

MARSHALL, C. C., et al. (1994, June). **Putting digital libraries to work**: issues from experience with community memories. Paper presented at the The First Annual Conference on the Theory and Practice of Digital Libraries, Texas - USA. Available at: <<http://csdl.tamu.edu/DL94/paper/marshall.html>>.

MEADOWS, J. (1997). **Changing patterns of communication and electronic publishing**. Paper presented at the IATUL Conference - "Scholarly Communication in Focus", Trondheim - Norway. Wikipedia. Floksonomy. Retrieved: 12/12/2006, 2006. Available at: <<http://en.wikipedia.org/wiki/Floksonomy>>.

Ana Alice Baptista

Miguel Ferreira

Dep. de Sistemas de Informação

Universidade do Minho

4810-058 Guimarães

Tel: +351 253 510 319

{analice, mferreira}@dsi.uminho.pt

Article received in: 2006, 12, 1

Article accepts in: 2006, 12, 1