

Sharing Open-Content Learning Resources in Emerging Disciplines

Darina DICHEVA¹, Christo DICHEV and Yi ZHU

*Winston-Salem State University
Winston-Salem, NC 27110, USA*

Abstract. This paper proposes a new type of collaborative learning repositories based on a hybrid organizational structure that leverages the potential of domain specific collaborative tagging in combination with conventional digital libraries classifications. It is exemplified in LinkedCourse - a learning repository prototype for rapid collaborative development, sharing and reuse of resources for emerging disciplines. The focus of the paper is on the collaborative semantic annotation and community formation, setting the socio-ontological framework of LinkedCourse. The implemented approach is aimed at keeping the flexibility of collaborative tagging and annotations within a domain-specific ontological framework.

Keywords. Participatory learning repositories, open content, collaborative semantic annotation, community formation

Introduction

Learning repositories were introduced as enablers for storing, sharing, reuse and repurposing of learning resources. However, there has been little sharing and reuse of educational materials through public repositories. One important aspect in the repository design that has been neglected is that the resources should be reusable and modifiable, keeping track of all contributors. Another factor limiting the widespread use of learning repositories is that they don't address adequately the specific needs of individual communities, which are typically formed around a shared domain of interest. Learning repositories are more likely to be successful if they are developed to meet genuine needs of a community. For example, in emerging disciplines the domain is evolving. The classification of the learning content (a form of a light-weight ontology), being domain dependent is also evolving. So is the domain vocabulary. This means that shared conceptualizations within not well bounded and evolving domains demonstrate a "work in progress" tendency. This suggests a hybrid classification framework that combines participatory processes and traditional classification approaches.

In parallel, we are witnessing a growing popularity of online communities that rely on mass participation and constant update strategies, such as social bookmarking. Many applications support building communities by empowering users to directly participate in a transparent collaborative process of content development. Typically

¹ Corresponding Author: Darina Dicheva, Winston-Salem State University, 601 S. Martin Luther King Jr. Drive, Winston Salem, NC 27110; E-mail: dichevad@wssu.edu.

they provide users with the ability to tag resources and to publish or share their tagging with other users. What makes the social tagging systems more attractive in repository context compared to conventional indexing approaches is that they support social interactions allowing users to connect to other users and to their resources and tags.

We believe that in a community of practice (a group of people connected by shared interests) tagging still has unexploited potential. A supporting hypothesis is that community-produced tags will be of higher semantic quality since they will reflect a vocabulary of community specific terms. This fact suggests creation of domain specific tagging mechanism able to leverage the implicit semantics emerging from the evolving tag structure and vocabulary. For example, community generated tags can be used as a source of terms to augment the evolving domain taxonomy, as they tend to represent the most current and natural domain terminology. The beauty of folksonomy is that the users do not have to learn any formal mechanisms; instead, they can tag content using freely chosen words. We believe that this basic freedom should not be sacrificed. Our approach is not to limit the tagging, but rather to set taggers in an appropriate context, namely, within a community formed around a particular course or topic.

Other challenges addressed in this work include stimulating participation and tracking content ownership. There are many incentives for publishing research publications in the academic community: academic reputation, promotion, institutional policy, etc. However, there are few such incentives to publish teaching materials unless in the form of an officially published textbook. This is one of the reasons why many lecturers tend to restrict the access to their materials to themselves or their close colleagues. The acceptance of learning content sharing on a community level requires adequate incentives implemented at a repository level. The learning repository discussed in this paper serves as a platform for content generation and reuse possibly by re-purposing and adapting of the original material. This presumes that a content unit can have many contributors, which leads to questions about ownership. We address this issue by using a Creative Commons License and keeping track of unit contributors.

The above considerations imply, that the new generation of learning repositories should depend on people's participation not only in content evolution but also in repository structure evolution and should address factors such as community formation and crediting authorship. To address existing needs related to emerging disciplines, we propose a community driven framework for rapid collaborative development, sharing and reuse of learning resources exemplified in *LinkedCourse* – a web application which we are currently developing.

The paper is organized as follows. We discuss the main ideas underpinning the proposed infrastructure in Section 1, introduce the infrastructure focusing on the socio-ontological aspects in Section 2, and present *LinkedCourse* in Section 3.

1. Towards Community-oriented Sharing of Learning Resources

Web 2.0 systems provide a medium for sharing and exchange of resources such as bookmarks, photos, videos, files, etc. Currently available “folksonomies” are geared toward casual social networking and prove successful for sharing and collaboration. Apparently, Web 2.0 offers a fresh approach that can be also used for sharing educational materials in emerging disciplines. It can foster the development of diverse communities of authors who are willing to share their material. Inspired by Web 2.0 and successful social bookmarking practices, we propose a novel community-oriented

framework for rapid knowledge exchange. We aim at infrastructure that supports participatory learning repositories, instead of the push models that traditional repositories provide, as users' participation can create content and keep it vibrant.

The key difference of our approach compared to the conventional learning repositories is that (1) it is focused on the connectivity of resources and users, and (2) it enables lawful modification of repository resources. Differently from social bookmarking systems that typically allow sharing links to someone else's resources, the goal here is (1) to support sharing of resources created by the participating members while addressing the corresponding intellectual property concerns, (2) users are expected to share not arbitrary bookmarks but links to learning content in a particular subject area, (3) the tagging process is based on a mix of controlled, semi-controlled and uncontrolled vocabularies (however taggers are still not limited in their choices).

Such an environment would be of interest to resource users who need learning material, possibly willing to endorse or critique the content and give credit to content providers, to resource providers who seek self-expression, community endorsement, and critique for the provided work and possibly attention from publishers and open-book supporters, to users who are interested in lawful reuse/modification of registered resources and/or seek collaboration for resource development, and to publishers looking for promising textbook authors based on demonstrated interest from the public.

1.1. Lawful modification of online learning materials

In emerging disciplines, courses are not well established and some can share multiple commonalities. Relatively minor modifications of some materials could effectively exploit the commonality of the bulk of shared content. The issue here is that there is no lawful way for instructors to modify learning materials found on the web even if they are willing to give proper credit to the authors. Thus, a mechanism for declaring that certain learning material is open and freely available for modification, extension, and reuse—as long as the authors are properly credited—is urgently needed.

1.2. Collaborative authoring

The availability of infrastructure, supporting the reuse of open licensed learning material in a subject domain, could make course content creation a much simpler task. Rather than writing a complete set of course materials, instructors can work on single topics, which are not covered and in which they feel experts. Authors can share their lecture notes, slides, assignments, problem sets, syllabi, reading lists, etc. They can form ad hoc working groups to collaboratively develop and adapt existing units. Through reuse of shareable units, a complete set of material for a specific course can evolve without waiting for the 'definitive' textbook to be published. This is especially important in emerging disciplines, where there are additional barriers for textbook writing: the initial market is relatively small and typically fragmented, and the lifetime of publications is often short (due to the rapid evolution of technologies).

1.3. Content contributors vs. content consumers

In a resource-sharing community, some members act more as contributors and others more as consumers. Various studies report that in participatory media (including wikis, photo-sharing sites, etc.), 5-10% of the users contribute half to all of the content [1].

Resource users however also play an important role through community filtering that ensures the promotion of good quality content. The contributor-consumer interaction offers richer opportunities though. If someone finds an open-content learning resource that is not an exact match of what they need, the potential places to look for a better match is in its “consumer” resources or in resources for which it plays a role of a “consumer.” Indeed, the content of a given resource might not match a particular instructional goal, however, someone may have adapted it appropriately. Therefore, we exploit a *resource connectivity*-based strategy for exploration.

2. *LinkedCourse* Framework

The proposed framework aims to support rapid, community-based development and sharing of learning resources while acknowledging and preserving the copyright of the authors. The keystones of the proposed framework are presented below.

2.1. Distributed content and intellectual property

The learning material registered in the repository is distributed and resides on authors’ websites. The repository contains only records with metadata for the original resources and their authors.

The registered resources are licensed under the Creative Commons license [2] that allows the content to be copied and redistributed, with or without modifying, and used for commercial or noncommercial purposes, provided the authors receive attribution throughout the use of the module, even when modified. This promotes the greatest possible sharing of materials.

2.2. Collaborative semantic annotation

The advantages and disadvantages of ontologies and folksonomies are well known [3, 4]. Ontologies can make content well organized, but they require time and expertise. Studies have shown that there is an ongoing reluctance among both users and institutions to create ontologies and metadata [5]. On the other hand, user-generated folksonomies can be more relevant and inspire discovery, but users lack discipline and expertise. Controlled tagging brings discipline but can create a gap between resource providers and users of learning collections, making the retrieval process tricky [6]. Yet, the existing approaches of combining ontologies and folksonomies have not demonstrated convincing results (see for example [4, 7, 8]).

Our approach is not based on simple coexistence of folksonomies and taxonomies as two different and complementary approaches for semantic annotation; instead, the idea is to mix them in an approach that lies somewhere in the middle. The suggested approach for sharing learning content is an attempt to combine some aspects from both worlds: conventional digital libraries and ad hoc classification. It is based on two observations:

- Based on their experience with personal folders, instructors are used to classify their material under courses, and subdivide it by course topics.
- Tags are inseparable from the context of the community in which they are

created and used [9].

Based on the first observation, the learning resources in our repository are divided into course collections. The course structure is employed not only as a predetermined classification framework but also to narrow down the user base to a particular community and thus to limit their tagging vocabulary by limiting the domain vocabulary as a source of tag choices.

We view tagging in three dimensions:

- as a process (from the viewpoint of the user's choice of terminology),
- as folksonomy (from the viewpoint of the generated collective vocabularies and resulting knowledge organization), and
- as a social activity (from the viewpoint of the social context of interactions and the resulting formation of communities).

The learning material resides in course collections. This is the place where storing, tagging and searching resources takes place. Thus courses are used as both an upper organizational infrastructure of learning resources and social infrastructure for user interactions and forming course level communities. Tagging in such communities will generate a conceptual structure as perceived by the corresponding community.

The *LinkedCourse* platform is aimed at aggregating community generated tags within semi-controlled vocabularies, metadata and domain-specific ontologies. The challenge here is in striking a balance between the open user-generated tags and the semi-controlled vocabulary. Our strategy is to constrain not the tag choices but instead the user base through limiting the domain, which serves as a common point of interest.

We are also extending the folksonomy model with upper ontologies, including WordNet, Dublin Core, and FOAF, in order to ensure some reusability and interoperability of information. More specifically, we use Dublin Core for annotating resources and FOAF for presenting authors' profiles. Entering complete Dublin Core data is optional. A minimal DC subset - the title, the author and the descriptions are derived from the resource submission process. FOAF information is obtained in two ways: retrieved by the FOAF RDF file of an user if they submit their FOAF URL, or generated from personal data entered by the user.

We envisage the use of three semi-controlled vocabularies. The first one comes from course names. Though course names put some boundaries on the tags variations, the choice of the descriptive terms is left open. The purpose here is to capture the domain specific terminology clustered around related concepts. For example, similar learning content can be found in courses named "Internet Systems", "Internet Technology", "Web Programming", "Web Design", "Scripting Languages", etc.

Another semi-controlled vocabulary comes from the resource types, e.g., lecture notes, code examples, assignments, free software, test samples, problem sets and solutions, syllabi, reading lists, etc. A third source of "controlled classification" comes from the automatic tagging of resources with contributors' information, e.g., username, institution, home page, etc. We plan to use *LinkedCourses* as an experimental environment for examining our hypothesis that collaborative tagging converges to controlled vocabularies that can be used as sources of terms for augmenting the evolving taxonomies of emerging domains.

The uncontrolled part of the tagging leaves users the freedom to pick arbitrary

categories for classifying learning resources besides the course and resource type classification. Such a feature will enable users to group resources by additional properties, including content-related, instructional, presentational, etc. Instead of having users haphazardly entering in tags to describe the resources they bookmark, *LinkedCourse* suggests tags used by the members of the corresponding community. In addition, using a tag cloud as a categorization system allows visualizing the power of the ad hoc classification. The tag cloud will allow users to navigate the collection by all properties used for grouping the resources and to discover interrelationships between groups that may not be apparent when navigating through courses.

2.3. *Community building*

The core idea driving *LinkedCourse* architecture was to build communities of instructors through collaboration and social tagging. One of our goals is to explore the feasibility and the potential of supporting the creation of sustainable communities of practice. In this aspect we aim at creating repositories that provide platform for discovering not only resources but also people. In contrast to traditional repositories, we provide a richer view on resources enabling users to see how they are used and who interacts with them.

The framework enables users to bookmark and vote on the quality of registered resources, to subscribe to receive information (through RSS and Atom presentations) when a new resource in a particular course, of a particular type, from a particular author, or tagged with a particular tag, is registered or updated in the system, etc. Several strategies are used to create incentives for reuse, including measuring and rewarding the contribution and use of content, combined with technical support that facilitates and encourages reuse. Besides, the framework supports the connection of users to encourage networking and help with further collaboration (for example, based on individuals' bookmarks). It also enables users' involvement in maintaining the website, implements a reward policy to encourage members' 'housekeeping' work and uses it to rank the involvement of community members, etc.

3. *LinkedCourse* Implementation

LinkdCourse design requirements include support for registering and tagging of learning resources, maintaining references between resources, users' reviewing and ranking resources, 'housekeeping' for maintaining good structure and content, intuitive navigation and searching throughout the collection of resources for finding courses or resources of interest, or other users with similar interests, provision of personalized resource spaces, and community building and communication.

3.1. *Services*

To enable this functionality we are implementing a service-oriented architecture. The main envisioned services are listed below.

- Registering resources: for each registered resource only a *resource entry* is maintained containing information, such as name, type, description, URL, etc.

- Tagging, reviewing, commenting, or voting for resources.
- Exploring resources: a combination of browsing paradigms are envisaged to support the exploration of resources: *facet-based browsing*, providing a five-dimensional view on the content (based on facets); *pivot browsing* [10], providing a lightweight mechanism to navigate an aggregated collection, *attribution & credit reference map*, and *tag clouds*.
- Community building and communication: services for support communication and collaboration between registered users, such as forums for discussing courses, resources, and tags and RSS and Atom presentations related to courses, resources, tags and authors; services to help involving new members and contacting existing members; statistics on the blogs' use, services for importing and exporting bookmarks, services for maintaining tags (e.g. reporting overlapping tags (having similar names), non-used tags, resources/tags with little metadata, or such that members voted as low quality/not useful).

3.2. Interface

Currently, the *LinkedCourse* interface contains two main spaces: global and private space (see Fig. 1). The private space contains the following tabs: *My Courses*, *My Resources*, *My Bookmarks*, *My Community*, and *My News*. The global space is the space where users can browse all information submitted to *LinkedCourse*: courses, resources, people, and tags. This is the space to which unregistered users have access. The private space is envisaged as a projection of the global space on a particular user. Therefore, it contains all courses, resources, and tags created or bookmarked by that user.

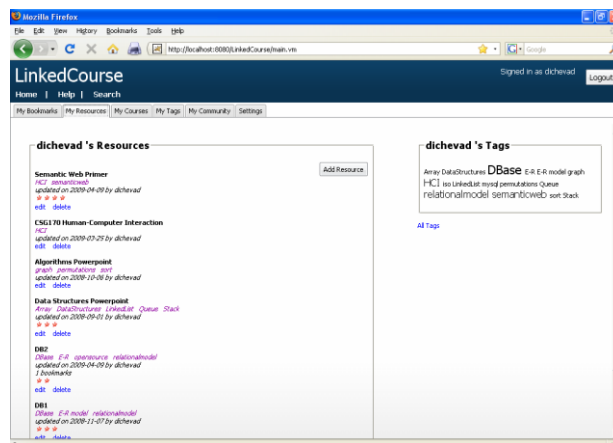


Figure 1. Screenshot of the *LinkedCourse* interface: a user's private space.

My Community connects a user to other *LinkedCourse* users. Each user can add registered users to his/her community for easy access to their learning resources/bookmarks/tags, as well as for more convenient contact to them through

special services.

My News is the space where a user can receive information for newly added or modified resources of their interest (through RSS feeds) or such exported to them by members of *their community*.

Users can subscribe to courses, tags, and people in order to receive information about new resources submitted to a specific course, by a specific author, or tagged with a specific tag.

4. Conclusion

In this paper we propose a framework for rapid collaborative development and sharing of learning resources for emerging disciplines, which is built on a set of intuitions shared by a wide range of academics: that knowledge should be open to use and reuse; that collaboration should be easier; that people should get credit and kudos for contributing to education research; that there should be a way for instructors to publicly acknowledge reuse of open content; and that the ability of authors and instructors to readily and dynamically access and update learning material is especially important in rapidly changing fields.

A professional community will succeed if the participating members perceive some value in their participation. In this case, the value is in the content that no single instructor is normally able to develop on their own. A pool of up-to-date teaching materials made available to community members through sharing and collaboration provides value and motivation for sustainability. Providing an audience and means for expressing the self is another value factor for contributors seeking reassurance. We believe that an appropriate infrastructure can turn a learning repository into a space where content attracts people and people bring others who use and further evolve it.

References

- [1] What makes people to participate? Online marketing strategies and techniques from Asia Pacific, (2007), available at: <http://www.rockyfp.com/blog/what-motivates-people-to-participate/>.
- [2] Creative Commons <http://creativecommons.org/>
- [3] H. Al-Khalifa, H. Davis: FolksAnnotation: A Semantic Metadata Tool for Annotating Learning Resources Using Folksonomies and Domain Ontologies. In: Proceedings of the Second International IEEE Conference on Innovations in Information Technology, Dubai, UAE (2006).
- [4] S. Bateman, C. Brooks, G. McCalla: Collaborative tagging approaches for ontological metadata in adaptive e-learning systems, In Proceedings of 4th Int'l Workshop on Semantic Web for E-Learning - SWEL'06 (2006).
- [5] J. Casey, J. Proven, D. Dripps: Geronimo's Cadillac: Lessons for Learning Object Repositories. 7 (2005).
- [6] Report on the Workshop of Learning Object Repositories as Digital Libraries, D-Lib Magazine, October 2006, available at: <http://www.dlib.org/dlib/october06/artacho/10artacho.html>
- [7] R. Lachica, D. Karabeg: Metadata creation in socio-semantic tagging systems: Towards holistic knowledge creation and interchange. In L.M. Garshol and L. Maicher (Ed.): Scaling Topic Maps, Topic Maps Research and Applications (TMRA 2007), LNCS, Springer Verlag (2007).
- [8] L. Specia, E. Motta.: Integrating Folksonomies with the Semantic Web, Proceedings of the 4th European Conference on the Semantic Web (ESWC 2007), Innsbruck, Austria, June 3-7, 2007, LNCS Springer 4519 (2007), 624-639.
- [9] P. Mika: Ontologies are us: A unified model of social networks and semantics, J. Web Semantics 5 (1) (2007), 5-15.
- [10] D. Millen, J. Feinberg, and B. Kerr: Dogear: Social Bookmaking in the Enterprise, In Proceedings of CHI 2006, Montreal, Canada, April 2006 (2006).