Information Architecture and Design Solutions Scaffolding Authoring of Open Educational Resources

Teemu Leinonen, Jukka Purma, Hans Pöldoja, and Tarmo Toikkanen

Abstract—This paper presents the open learning object repository and collaborative authoring platform LeMill (http://lemill.net), which has over 7,500 members and over 8,500 reusable learning resources (situation in October 30th, 2009), all created by the community members. The design of LeMill has tackled numerous challenges that hinder the authoring and sharing of educational resources by communities of teachers. This paper describes the research-based design process that was used to solve these challenges. The information architecture of LeMill scaffolds authors toward collaboration and sharing. The licensing scheme encourages reusing and remixing of educational content. In order to make LeMill easy to learn and use we have avoided technical terminology and complicated metadata forms in the user interface. As an open community we have also tackled multicultural and multilingual issues. In this paper, we present the information architecture and design of LeMill, including the technical solutions. We believe that our design solutions will contribute to the goal of creating an open educational resource ecosystem.

Index Terms—Computers and education, social learning techniques, social networking, user-centered design, user generated learning.

1 INTRODUCTION

There is something deceptively simple in open educational resources (OER). Almost anything can be used as an educational resource [1] and anything that is offered for free and without major social or economic expectations can be understood to be open. It can look like the internet is full of open educational resources, ready for teachers to adopt and to use. In reality, this doesn’t seem to happen. To get people to use open educational resources, there have been projects to define, package, and share them. Currently, research in educational technology has proposed several definitions on what OERs are [2], [3], there are some generally shared standards on how to represent them, and numerous repositories for collecting and sharing them.

As far as we can see, there are still a few missing links before the adoption of OERs can take place in everyday teaching and learning. The most important missing link is that there is no room for OERs in the everyday activities of a median teacher. To fix this, we would need to fix the daily activities of teachers, and we cannot do that directly. However, what we can do is design tools that will support new kinds of everyday practices of teachers and learners.

Because we cannot simply push OERs and the changes they necessitate on teachers and learners, we should try to minimise the required conceptual shifts and changes in learning activities. LeMill (Learning Mill) is a Web community for finding, authoring, and sharing educational resources, designed for easy integration with teachers’ existing meaningful tasks and needs. LeMill was developed in 2005-2008 within an EU funded CALIBRATE project (http://calibrate.eun.org), with an initial premise of providing a toolbox for collaborative authoring of learning resources. After CALIBRATE, further development and dissemination has been done in the contexts of the Estonian Tiger Leap Foundation and the EU funded Finnish AVO project (Open Networks for Learning, 2008-2011).

There are implicit and explicit assumptions about what good OERs are: They should be relevant to the learner and thus easily modified to fit the learner’s needs. They should be of good quality and contain no factual errors. They should disclose their point of view and in the case of science be free from bias. They should not have hidden costs or prohibiting limitations on use. A good learning resource should also be able to “travel well,” to be easily translated and recontextualised [4].

We believe that these requirements can be met by having the resources edited collaboratively and freely online with no restrictions on participation. If OERs can be freely edited, they can be customised for the needs of individual teachers and learners. If they have multiple editors working on them, versions will evolve, making errors and biases easier to find and correct. The license scheme must permit all this: free editing, sharing of edited versions, and combining versions to form new resources. These requirements point toward a uniform and nonrestrictive licensing scheme for all resources.

Since we want LeMill to have good OERs and collaborative authoring seems to be the way to do it, but teachers’ existing activities, tasks, and skills do not necessarily include collaborative authoring, we designed...
LeMill to provide scaffolding for teachers, so that what they do with their existing skills becomes directed toward collaborative creation.

The main research question of this paper is: How can a Web service design promote use and creation of OERs?

This paper will start by presenting the general challenges related to the design of open educational resource systems, which are not only technological but also social systems. The paper continues by defining relatively concrete design challenges that are part of the overall “wicked problem” landscape. Wicked problem is a concept used by Rittel [5] when referring to the nature of problems common in planning and design practice. After presenting the design challenges we introduce the design methodology used in the process. We call the methodology a research-based design process with focus on software as hypothesis [6]. We continue by presenting design solutions that scaffold authoring of open educational resources. The solutions are mainly decisions related to information architecture: ways of organising, structuring, and enabling collaborative authoring and sharing of educational resources online. We conclude by comparing LeMill’s design solutions to other popular repositories and learning resource authoring services.

2 Design Challenges

When starting the LeMill project our initial design challenge was the assumption that European teachers do not share their learning materials nor do they improve them in a collaborative way. Beyond this challenge we can see more general challenges related to European and international educational politics. These are, for instance, differences in the results of educational systems in different countries, which cause problems in recognising educational degrees in other countries. Not recognising educational degrees in a pan-European level hinders the free movement of people, which is one of the basic components of the European Union and acknowledged as a fundamental right for EU citizens.

While our task was not to solve the problems related to the rights of EU citizens, as designers we considered it important that we are aware of the the big picture. Focusing is not possible if one does not know the context. When narrowing down the design challenge to problems related to sharing of learning materials and improving them in a collaborative way, we already implicitly defined the general design solution. This is common in design thinking where the fact that problems are wickedly incomplete and often contradictory is taken for granted [7], [5], [8]. A designer’s way of approaching a problem includes the idea that all problems have multiple solutions and every formulation of a problem is simultaneously an attempt to solve it. According to Nelson and Stolterman, ordinary problem solving is reactive to unwanted states, while designing is about creating a positive addition to the present state [7]. The designer cannot assume that the truth about optimal design is there to be found. Instead, the designer can point a way and say that choosing this way has some benefits.

We chose to work with teachers because they are the part of each country’s educational system that has direct effects on learning outcomes. Within the educational system, changes in teaching are easier to track than changes in learning. If we had worked outside of the educational system (self-learning, open learning, and networked learning) we would have positioned ourselves as outsiders to our pan-European problem, that of recognising educational outcomes from different educational systems.

During design and development the main design challenge broke down into smaller, often more urgent subchallenges. These subchallenges represent recurring themes in LeMill’s design and we think they can be expressed as general design problems in OER repositories and services. These subchallenges are:

1. Lack of collaboration and peer production of learning materials.
2. Lack of reuse and remixing.
3. Limited access and poor usability.
4. Barriers related to multilingualism.
5. Poor use of the underlying principles of the Web, such as openness and “linkedness.”

In Section 4, we will argue for and explain our solutions to these subchallenges.

3 Methodology and Design Process

Our design methodology is called “research-based design with software as hypothesis” [6]. In several earlier design and research projects aiming to develop new learning technology (FLE, Fle2, Fle3, MobilED, Hauki, and Kuha) we have noticed that people create meaningful ways of using the tools that surround them, and, from the perspective of tool design, often do not know beforehand what tools they really need. The consequences and the affordances of the tools are realized only when they are used in the real world. With LeMill, our aim was to design learning technology in an open dialogue between designers and the target group (in this case teachers) and provide them with software prototypes. With these prototypes we can design affordances (as understood by Norman [9]) that will likely make sense for the teachers at first glimpse.

The relationship between meaning created in action and tools can be illustrated with an example from school architecture. An auditorium and a teacher’s podium are tools that form learning spaces. The architecture, fixtures, furniture, and props in the space quite openly communicate and support certain types of teaching and learning. In complex social activity systems, all new tools bring changes to existing activity systems. A new tool should communicate the changes needed in the system. In our context of educational technology this means that the designed tools and artifacts are always also communicating what teaching and learning with them could be like. In this sense, software tools can be presented as hypotheses about teachers’ activities. They can succeed or fail at inducing activities that teachers are willing to integrate into their everyday teaching.

Research-based design with software as hypothesis is not to be confused with design-based research. In design-based research [10], [11], the aim is to do research with designed interventions into real-world situations [6]. In design-based research design interventions are a research method. In research-based design, the design is the main outcome and anthropological (or quasi-anthropological) research helps to
draw routes to that outcome.\footnote{1} Our process of research-based design aiming to design a new tool is divided into four iterative phases which happen partly in parallel:

1. contextual inquiry,
2. participatory design,
3. product design, and
4. production of software as hypothesis (see Fig. 1).

The process resembles a hermeneutic circle where all research and design operations increase the researchers’ and the designers’ understanding of each other and the context [6].

As part of the research-based design process, LeMill was developed by using the principles of scenario based design [12] and agile software development methods [13]. The design process was carried out and documented in a publicly available software development environment, called Trac (http://lemill.org).

The contextual inquiry phase included observation of teachers’ computer use with a special focus on searching of online learning materials. In addition we benchmarked several other online services. These were MediaWiki (http://www.mediawiki.org), MIT’s Open Courseware (http://ocw.mit.edu), Connexions, (http://cnx.org), MERLOT (http://www.merlot.org), Pachyderm (http://www.nmc.org/pachyderm), and eduCommons (http://cosl.usu.edu/projects/educommons). Teachers were also asked to use these systems and relate their experiences.

In the contextual inquiry phase our impression was that median teachers in the year 2005 hardly used online learning materials. According to a study conducted in 2006 in European countries [14], 40-85 percent (depending on country) of teachers have used computers in class in the last 12 months, 70-90 percent consider themselves competent in using ICT, and 70-95 percent have used material retrieved from the internet. However, in our observations, most of the teachers used very limited computing skills to produce learning resources: a basic knowledge of office software and using copy-paste to add internet resources was enough for most of the teachers.

Participatory design sessions with one researcher-designer and 2-3 teachers were organized in Estonia, Finland, Hungary, and Norway. The teachers read prepared scenarios and then discussed each scenario in a structured group interview led by the researcher-designer. The participants were then asked to visualize the proposed system as they imagined it and explain their drawing. The researchers recorded the sessions and wrote summaries of them into the Trac system.

During product design we realized that we were designing and developing a new tool: something to which none of the existing learning resource authoring tools could contribute. When analyzing the participatory design sessions, we noticed that teachers often do not know what kind of tools they really need and their wishes are influenced by tools that they currently use. We had to balance this reliance on existing tools and their frameworks with the affordances of the new tool. This balancing was an iterative process with teachers, as we gradually became better at understanding the perceived usefulness of each feature.

Our software as hypothesis is the LeMill service. We believe that with it—a simple Web-based tool that provides a clear structure for learning resources—teachers can create communities of practice that share and create open educational resources.

4 Design Solutions

This section presents our solutions to the design challenges outlined in Section 2.

4.1 Scaffolding Collaboration and Peer Production

The lack of collaboration and peer production of learning materials was the first of the design challenges. The issue
was considered from several points of view, including the basic structure of the service, the level of “wikiness” that would rather benefit than become an obstacle for collaboration, basic concepts and their relations, and the workflows of collaboration, and peer production.

The overall tasks that teachers should be doing in LeMill were deduced from participatory design sessions (http://lemill.org/trac/wiki/DesignSessionResults) and later from workshops with teachers. The structure of LeMill and the available features on each page should provide scaffolding for these tasks. The main tasks were defined as 1) finding resources to use (illustrations and exercises), 2) finding new teaching methods, and 3) modifying resources to better suit a particular learning context. In the participatory design sessions we noticed that when planning their teaching, teachers are primarily interested in using some new teaching or learning methods or tools and only secondarily are interested in the curriculum in which these could be utilized. Based on this discovery, we decided that the top level navigation in LeMill should be split into four segments: Content, Methods, Tools, and Community.

Also apparent from the design sessions was that teachers are generally wary of collaboration and resource creation, pointing out problems with copyright, motivation, and high threshold for joining a new community. While from a workflow perspective LeMill resembles a wiki, its user interface is quite far removed from that of Wikipedia for instance. While a traditional wiki-like interface promotes coediting among a community of technologically savvy users, this is not sufficient to encourage teachers to collaborate, as participation in wikis requires mastering a relatively complex syntax and the environment requires joint continuous efforts to maintain a coherent structure. To support collaboration and peer-production among teachers we noticed that on top of the “wikiness” we must have additional scaffolding that presents the information architecture and the concepts in a language that teachers already know and are comfortable with.

For instance, Methods and Tools are familiar concepts for pedagogically minded teachers. Methods are descriptions of various pedagogical methods, activities, games, and other ways of teaching and learning. Methods also represent LeMill’s unique take on Learning Objects (LOs) and Learning Design (LD). While most online LO repositories primarily contain learning objects for learners, LeMill focuses on resources that teachers can utilize to improve their teaching. LOs have been criticized for their unfounded promise of Lego-like combinatorics, which would only be possible if they were instructionally empty [15]. Learning Design (LD, [16]) on the other hand adds instruction theory to learning objects, but the level of description required for IMS-LD modeling is too cumbersome for our needs of easy access and online editing, and on the other hand is incapable of representing some advanced pedagogical models [17]. In LeMill methods are treated like other learning resources and it is up to the teachers to decide how a certain method should be used with certain content. LeMill’s LOs are not supposed to be fully machine readable or used in automated instructional sequences. It is assumed that there is always a teacher to decide how resources are to be used. If a teacher is there to contextualize the resources, the decontextualized nature of LOs is not a problem [18].

The Content section contains more typical learning resources. Content resources are built on one of six available templates. The templates are basic scaffolding tools that make it easier for people to create Web content [19], [20]. Our templates are: webpage, presentation, exercise, lesson plan, school project, and PILOT (Progressive Inquiry Learning Object Template [21]). The main concepts and divisions of LeMill are described in Fig. 2.

An important aspect of LeMill is the authoring workflow. We would have preferred to keep the authoring workflow as simple as possible, but participatory design sessions revealed that teachers had many reservations about releasing unfinished or partial resources. So there had to be a division to drafts and published content. However, that division does not need to apply to resource types that don’t have a precedent for such division. Our judgement has been that methods, tools, references, PDFs, learning stories, and media pieces do not need to have a draft version. Only template-based content types should have drafts (see Fig. 2).

Drafts were initially created as public to encourage collaboration but without prominent author information to lower the threshold of creating unfinished content. Because of feedback of teachers we had to make drafts private by default and made a third option of “public draft.” Publishing a resource is encouraged by allowing only published resources to have cover images. When resources are published, they are visible for all and editable by all.

We initially designed the communities in LeMill to form around collaborative learning resource creation. Each resource can be adopted by a group. Anybody can join a group, but joining a group is the precondition for editing a resource. Later, when we noticed that groups were also used to form courses or workshops, and that these groups collected interesting learning resources, we allowed resources to be adopted by several groups. Discussion about learning resources was initially limited to happen within groups, but as these discussions were rare, the threshold to start one stayed high. We tried linking these discussions to resources with similar “discussion” links as seen in Wikipedia, but the concept became complicated when resources could belong to several groups. Finally we migrated all discussions about resources to happen within the resources themselves and having groups only aggregate these discussions.

We have also tried two solutions for branching different versions of a resource, but we are not satisfied with either. At first we allowed versioning, but found that teachers were too eager to use it in avoiding modifying each other’s works, with a detrimental effect to collaboration. The second attempt was when there was an existing biology textbook that an author wanted to publish in LeMill, but didn’t want it to be changed by anyone. We allowed the locking down of resources so that only author can edit them, but with the condition that there can be new branches. This feature also seems to be misused, and we are planning in removing it.

Authoring of learning resources and collaboration around them is encouraged by a teacher’s portfolio. For every teacher
it keeps track of where the teacher has participated in creation or editing of resources and aggregates these resources. The portfolio also has room for profile and personal information and interests, and these can be used to find other teachers with similar interests.

Individual teachers are also encouraged to collect interesting or high quality resources into their personal collections, which are visible to others. These collections can be formed around any theme, such as “interesting math resources,” “good pedagogical advice,” etc. A collection can also be used to create a lesson or course plan: by adding content, methods, and tools into a collection, a teacher can create a package that has much of the information that is needed to teach a lesson or a course (Fig. 3).

Teachers can also add “teaching and learning stories” to their collections. These are simple free form descriptions of a collection explaining how the teacher plans to use them or has used them in teaching. Resources that are used in teaching and learning stories automatically and prominently link back to them so as to provide examples and ideas on how to use them. This design addresses the common problem of learning objects not having contextual information about how they should be used [22], [18].

4.2 Scaffolding Reuse and Remixing

Reusability of learning resources has both technical and legal aspects one must consider when designing a service and a tool for this purpose. In the participatory design sessions we found out that teachers are aware of the copyright issues but many of them have a rather practical stand on them. The principle seems to be that if some online content is found useful in teaching and learning it can be used for this purpose freely, including copying, printing, remixing, distributing, etc. However, teachers perceived it as a plus if they could do so legally.

Traditional copyright laws give the creator of an original work the exclusive right to decide how their work is distributed and if it can be adapted. This is a major obstacle for the reuse of learning resources. Learning resources that teachers find from the Web may often need to be adapted to a certain learning context and target group. Doing this for personal learning purposes poses no problem, but distributing the adapted version is a
conundrum. According to copyright laws teachers need an agreement from the holder of the copyright before they can adapt and distribute the learning resource. These legal obstacles can be solved when learning resources are published under an open content license.

From the beginning of the project it was clear that all the content created in LeMill must be published under an open license. However, it was important to choose a licensing scheme that will both protect teachers and give them the possibility to reuse content created by other people. In 2005, when we were pondering this, several licenses were used for educational content. Some systems used GNU Free Documentation License (Wikipedia and other initiatives of the Wikimedia Foundation), some used Creative Commons licenses with noncommercial restriction (MIT OpenCourseWare) and some allowed users to choose between different Creative Commons licenses or all rights reserved (Flickr).

The first important decision was to use the same license for all resources that are created in LeMill. This enables teachers to remix all the resources that they find in LeMill without having to think about license compatibility issues. Second, we decided to choose one of the Creative Commons licenses because their licensing scheme is developed to be understandable by a wide audience. Finally, we were considering between Attribution-Noncommercial-Share Alike (BY-NC-SA) and Attribution-Share Alike (BY-SA) licenses. The noncommercial restriction limits the possibilities of reuse. In the educational context it is problematic for many meaningful ways of reusing content [23]. Therefore, we decided to choose the Creative Commons Attribution-Share Alike license for all content created in LeMill. Back in 2005, it was not the most popular choice for educational content, but recent developments in the field show that it was the right decision. The Wikimedia Foundation has migrated from GNU FDL license to Creative Commons BY-SA license [24]. This made it possible to remix Wikipedia (and other Wikimedia projects') content with LeMill content.

In addition to the new possibilities with Wikipedia and Wikimedia Commons, Connexions has all their content under a compatible CC Attribution (BY) license. The popular photo sharing site Flickr has millions of images under CC BY and BY-SA licenses. A growing number of content with licenses compatible with CC BY-SA makes
LeMill part of a larger OER ecosystem, while content that is under complete copyright can be used by linking or embedding under the Fair Use conventions. For example it is possible to embed videos from YouTube to learning resources in LeMill.

While there are no legal restrictions for remixing the content that is created in LeMill there are still some technical limitations in order to keep the system simple. With remixing we understand the combination of two or more learning resources. So far we have not developed special tools for combining parts of learning resources, but one content type—the media piece—is intended to be used with and within other resources. Web pages, exercises, presentations and PILOTs can all include media pieces. When adding a media piece to a resource, the author can search from existing media pieces or upload a new piece.

The remixing culture makes it easier for us to concentrate on our main focus by outsourcing some of the difficult parts of content creation to services specialized for them. For example, we have had trouble designing and implementing a fast and easy way to create presentations or slide shows as learning resources. Building them from media pieces is cumbersome and leads to dozens of uploaded slides with minimal reuse value. Teachers also want to upload existing Microsoft PowerPoint slides as learning resources, which leads to additional problems because then online editing and improving is not possible. Our current solution is to run OpenOffice.org as daemon to export PowerPoint slides into images once they are uploaded, and thus, create editable and “granular” presentations from uploaded material. In addition to putting presentation into LeMill users may use any of the external presentation hosting and creation services like SlideShare (http://www.slideshare.net) or 280 Slides (http://280slides.com) and embed slideshows from there into resources.

Embedding media from another site is actually another popular way to remix content in LeMill. The common method for embedding is to copy and paste an “embed code,” a piece of HTML, into a page or blog post. In workshops we have noticed that many teachers are accustomed to office software paradigm for creating content. In office software copying and pasting is the most common solution for moving pieces of text or images from one document to another. Thus we assume that copying an embed code from a site is the cleanest and most versatile way to embed content. The simplicity of copy and paste outweighs the additional user interface clutter that graphical remixing tools would add.

4.3 Scaffolding Access with Minimal Metadata

We suspect that the problem of limited access and poor usability of educational repositories in general is related to different perceptions of what is important for repository curators and repository users [25]. In the participatory design sessions we noticed that average teachers do not know what “metadata” is or see how it could be important to them. Nevertheless, teachers use different kinds of metadata in their daily work. From a technical perspective metadata is important to have, but for teachers it should be invisible, implicit or obviously useful [26].

LeMill is a repository of educational resources. Repositories store objects and metadata, and metadata is there to help find relevant data objects and communicate to other systems about their existence. There are several metadata schemes for educational resources. The Learning Resource Exchange (LRE) Metadata Application Profile v3.0 [27] of the IEEE Learning Object Metadata (LOM) standard [28] defines the metadata that European learning resource repositories should support. These standards can provide a solid base for designing an educational resource repository.

LeMill is built on Zope (http://www.zope.org), a transactional object database. Using an object database allowed us to be very flexible with the actual data model and start with a very minimalistic object scheme. Object schemata are easily updated to have new or changed fields. Having an architecture based on custom object types suggests using adapters to accommodate them to different metadata schemata instead of trying to keep the data structures themselves uniform and compatible. Educational resources from LeMill can be harvested with the Open Archive Initiative Protocol for Metadata Harvesting v2.0 (OAI-PMH) [29] as LRE LOM objects or using DublinCore metadata. In short, the actual data model is there to reflect the priorities of teachers creating content as far as we know them, while satisfying metadata harvesters and queries from other systems comes secondary and is done with adapters. Technically this has proved to be feasible and can be seen as a local mapping solution to problems of metadata interoperability [30].

In LeMill’s user interface we altogether avoid the word “metadata,” because teachers’ existing workflows for preparing material for classes do not use the term. Teachers have a very contextualized short-term need, whereas curators think about the general form and future accessibility of data [25]. If metadata is not perceived as essential for finding resources [25], [31], [32], then we suspect that adding such metadata to content will be perceived as an extraneous and unnecessary task.

To make some metadata relevant, we encourage teachers to browse LeMill. For example, the Content section’s front page has emphasized links to browse by language, subject area, target groups, and tags, with links to the three most popular tokens for each. Internally, all these browsing options are metadata categories. All of them except tags come naturally from teachers’ needs. Free form tagging is a concept familiar from social software and it has been found that teachers adopt it well [33]. After limiting results with one criterion, the browsing view allows the addition of other criteria from drop-down menus so that teachers can end up browsing for example resources in English that are about History and suitable for 10th grade students and have the tag media.

We assume that the usefulness of metadata in browsing encourages teachers to enter similar metadata to their resources. The data that LeMill collects that can be understood as metadata is presented in Table 1. Only the first four are explicitly asked from teachers and they are all optional. The rest of the fields are created automatically. The teachers may enter metadata when creating the resource, or they may complement them later. Completion—the “wiki-way”—can be done by any user. LeMill is integrating flexible community-based metadata creation to automated metadata gathering, as described by Duval [26].
When presenting metadata fields to teachers we have to use the same terminology as teachers do. For example, in workshops we found out that teachers prefer to use grades instead of a typical age range. Because of this we combined three elements from LRE LOM (“Educational.Intended End User Role,” “Educational.Learning Context,” and “Educational.Typical Age Range”) into a new element named “Target group.” Vocabulary values for this element include all the primary education grade levels, preschool education, higher education, adult education, special education, and teachers.

We have developed ranking algorithms for calculating scores for content, methods, and tools. The score will depend on the way people work on the resource and on the actions that other people have with it. Each object will get initial points when it is published. The score will rise when it is edited further, illustrations are added and external resources are embedded. Since our aim is to support collaborative authoring we will give more points when the resource is edited by more than one member. Points will be added to the score each time other people bookmark the resource into their collections. As a result, resources that are edited by several people and belong to several collections have a higher score. The scores are used to sort search results, generate tag clouds, and display featured resources on the section front pages. These algorithms are modified periodically as we try to balance results to both encourage collaboration and to reward individual efforts.

In a similar way we calculate scores for community members. The member score consists of three parts. First, we sum up the scores for all content, methods, and tools that the member has created. Then the social activities such as sharing teaching and learning stories, participating in the groups, and being added as a contact are scored. Finally the member will receive additional points for fully filling the member profile.

### 4.4 Scaffolding Multilingual Use

The participatory design sessions were carried out in four European countries—Estonia, Finland, Hungary, and Norway—in four different languages. From the very beginning it was clear that we were designing a multilingual and multicultural tool and service. The results should include a discussion page to give meaningful feedback about the resource. This approach is being used in LeMill, but conclusive results have not yet been gathered. One aspect of this approach is that it blurs the line between commenting and editing content. If you have a constructive comment on a resource, will you write it into the discussion page as a comment, or directly edit the resource itself to reflect the changes, or both?

We have developed ranking algorithms for calculating scores for content, methods, and tools. The score will depend on the way people work on the resource and on the actions that other people have with it. Each object will get initial points when it is published. The score will rise when it is edited further, illustrations are added and external resources are embedded. Since our aim is to support collaborative authoring we will give more points when the resource is edited by more than one member. Points will be added to the score each time other people bookmark the resource into their collections. As a result, resources that are edited by several people and belong to several collections have a higher score. The scores are used to sort search results, generate tag clouds, and display featured resources on the section front pages. These algorithms are modified periodically as we try to balance results to both encourage collaboration and to reward individual efforts.

In a similar way we calculate scores for community members. The member score consists of three parts. First, we sum up the scores for all content, methods, and tools that the member has created. Then the social activities such as sharing teaching and learning stories, participating in the groups, and being added as a contact are scored. Finally the member will receive additional points for fully filling the member profile.

## TABLE 1

<table>
<thead>
<tr>
<th>Field name</th>
<th>Teacher’s input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tags</td>
<td>free text</td>
</tr>
<tr>
<td>Subject area</td>
<td>multiple selection</td>
</tr>
<tr>
<td>Target group</td>
<td>multiple selection</td>
</tr>
<tr>
<td>Language</td>
<td>suggestion based on the teacher profile’s languages, single selection</td>
</tr>
<tr>
<td>Publication status</td>
<td>for most resources ‘published’ or ‘deleted’, for Content types also ‘draft’ or ‘private’; altered by actions ‘publish’, ‘delete’, ‘undelete’ or ‘retract’ shown when applicable</td>
</tr>
<tr>
<td>Cover image</td>
<td>asked when publishing a resource, can be changed later; for media pieces that are images, automatically use thumbnail image</td>
</tr>
<tr>
<td>Creators</td>
<td>automatically added as creator, order of authors depend on size of contribution</td>
</tr>
<tr>
<td>Rights</td>
<td>CC BY-SA automatically for most; with references and media pieces several options that refer to original license</td>
</tr>
<tr>
<td>CreationDate</td>
<td>automatically added</td>
</tr>
<tr>
<td>ModificationDate</td>
<td>automatically updated</td>
</tr>
<tr>
<td>Id</td>
<td>automatically generated from resource’s title</td>
</tr>
</tbody>
</table>
started with everything in the same pool, and gradually added functionality to allow for different languages to separate to their own resource pools. The upside of keeping everything together is that the repository doesn’t appear empty to a representative of a minor language, while the downside is that search results may be flooded with resources in a language that the teacher doesn’t understand.

From the beginning each teacher was able to define in their profile the languages that they are fluent in, in the order they think is most suitable, usually placing their native language first, followed by other languages that they can use. This became a very important tool as we noticed that LeMill was starting to be dominated by a few small languages.

Nearly all resources in LeMill have a specified language, and those that don’t are causing problems, so in the future they will need to be tagged with a language as well. We use teachers’ profile language information to customize both the user interface and the listed resources in search results, featured resources, and browsing views. The list of matching resources is sorted by languages, and then by the individual resource’s popularity score. Thus, in search and browsing results, teachers will first see matching entries in their native language, in popularity order, followed by resources in their secondary languages, in popularity order, language by language, and finally in English, if English was not already included.

Resources in languages that the teacher has not listed in their profile will not be shown at all unless explicitly searched for. Teachers can of course access them if they find them. This feature acts to form language clusters within the repository, as the language skills of the teachers define the borders surrounding clusters. Multilingual teachers will of course be positioned as mediators between language clusters.

LeMill encourages translation of resources. Each resource has a link to translate it, and the resource has links to already existing translations. Translations are not assumed to be identical, and certainly cannot be, as the original and the translation can both be further developed by other teachers. The intention is to facilitate the spread of good resources and teaching ideas.

We’ve identified some specific problems regarding resource languages. Images don’t usually have any language content, but their descriptions and titles are written in some language. Would it make sense to tag an image with the language of its description? While it is informative, teachers could easily reuse images regardless of their description language.

Another problem concerns collections, which can contain resources in multiple languages, in addition to the collection’s title and its own description. What should the language of a collection be, if its title and description are in Estonian, but all or most resources are in English?

A third problem concerns resources related to language studies. If a resource contains text in English, and instructions in Estonian, which language should it be tagged with? English teachers in Estonia will most likely try to find material for their courses by looking into the pool of English resources, but having a resource that is partly in Estonian will be quite problematic for English teachers in other countries. Short of having separate metadata fields for “teaching language” and “content language,” this issue is still unresolved, partly because there is a similar problem with referencing to resources outside LeMill. The referenced resource can be in a different language than the actual reference description and explanation, but there is an ambiguity about what the language field is referring to. LRE LOM’s approach of asking for language in “General.Language,” “Meta-metadata.Language,” and “Educational.Language” allows all of these specifications. However, in most of the cases selecting the language of a resource is obvious, and having these three fields would feel like an unnecessary complication.

A fourth problem is the issue of tags and their languages. Currently tags carry no language information, which means that tag clouds are quite multilingual. It would be technically possible for us to convert all our tags to tuples of tag and language code, but keeping the process of adding tags simple would then be a formidable design challenge. Tag language could be inferred from the language of the resource, except that many content resources are about foreign language learning and thus contain two languages and are labeled in differing ways (see previous paragraph). Inferring tag language from the user’s profile would require each tag to have a language specified separately, as tags can be edited by anyone and adding a few tags in your own language shouldn’t change the language of other tags. Even a dictionary-based solution has its limitations, as some words can occur in multiple languages and mean either the same thing, or different things.

A multilingual site attracts users from multiple countries and cultures. At the time of writing this paper, LeMill has teachers from 56 countries. Each group of sufficient size seems to form its own codes of conduct and practices of using LeMill. Teachers in some countries have started to write individual course plans as method descriptions, while others have written them into the content section. In some countries teachers need to be able to present their lesson plans and show how they connect to accepted learning goals, whereas in other countries there is no need for such detailed plans. As a response new content types were added to the content section: lesson plans and school project plans. One example of cultural differences is the popularity of history as a subject area in Georgian resources, whereas in other languages it is one of the least used categories. A more detailed comparison of community practices is difficult because of language barriers, which also forces us to trust in community self-organization and self-policing.

### 4.5 Scaffolding Creation of Small Pieces Loosely Joined

Our fifth design challenge was the poor use of the underlying principles of the Web, such as openness and “linkedness.” What we mean with openness is that anyone can join the system, create new resources, have them link to any resource anywhere, and to link to these resources from anywhere on the Web. Because of its elective, haphazard, and unlimited fashion of linking, the Web is said to be formed from small pieces loosely joined [34]. A powerful side effect of ideological linkedness is that because search engines index resources by following links, resources in highly linked and openly traversable repositories are very visible in general search results.
As noted before, our approach to standards is not to build from standards, but to build from teachers’ needs and have the result adapt to standards when necessary. All textual resources in LeMill are presented as XHTML. Multimedia uses the normal Web-acceptable image formats, and some rich media are displayed using Flash, which can be considered a de facto standard. As previously discussed, we considered IMS Learning Design (IMS LD) [16] to provide an interesting and advanced description language for pedagogy, but we found it too complicated for easy access and also that it had important restrictions in covering dynamic group behaviors and other advanced iterative methods [17].

If the resources in a repository cannot be exported and transferred to other infrastructures, it cannot claim to be truly open. The most popular LO transfer format is SCORM, and collections from LeMill can be exported as SCORM sequences, stand-alone Web file packages, or pdf booklets. These allow teachers to export their collections and set them up on a Web server, import them to any SCORM-compliant LMS, or to print them as handouts.

Refractories or metarepositories that only store the metadata of resources have lately been quite popular, and any open repository should acknowledge them and provide access to them. OAI-PMH [29] support was built into LeMill, as it seemed to be the leading protocol in querying repository contents, had sufficient support for LO metadata, and was already supported by other platforms. LeMill also provides RSS feeds for search results and other dynamic pages. A summary of currently supported protocols and views is presented in Table 2.

<table>
<thead>
<tr>
<th>Reader</th>
<th>Provided protocols and views</th>
</tr>
</thead>
<tbody>
<tr>
<td>teacher</td>
<td>web page, ‘student view’, zipped offline web pages, SCORM-package, PDF</td>
</tr>
<tr>
<td>RSS-reader</td>
<td>RSS 2.0</td>
</tr>
<tr>
<td>metarepository</td>
<td>OAI-PMH, supported formats LRE, LOM and DublinCore</td>
</tr>
</tbody>
</table>

LeMill’s main design goal is to foster peer production, LeMill omits a few common OER repository features in order to make participation and co-creation easier. LeMill differs from major repositories like Connexions (http://cnx.org), MERLOT (http://www.merlot.org), OER Commons (http://www.oercommons.org), and its sister project Learning Resource Exchange for Schools (http://lreforschools.eun.org) by keeping the user interface much more simple. The resources cannot be rated and they have less visible metadata.

Simplicity in browsing makes simplicity in editing easier to achieve. In Connexions, creating Modules is done in their own CNXML language; in MERLOT, OERCommons, and LREforSchools, resources are submitted by providing a link and entering metadata, after which the resource goes through a review process. Even as resources are often published with open licenses, only the author or editor can modify the resource. Only wiki-based OER projects like Wikiversity (http://wikiversity.org), Wikieducator (http://wikieducator.org), and LeMill expect collaborative editing to be the default.

Mixing languages is also typical for LeMill, but as a sitewide feature it may be a passing phase, as content in foreign languages are only displayed when enough resources in familiar languages are not found. MERLOT and Connexions use English as the only user interface language. Connexions has ≈90 percent of resources in English and MERLOT’s ratio is unknown, although probably in the same range. Wikiversity and Wikieducator use separate subsites for every language, with courses linking to other languages if the translations are available. There is also a multilingual beta.wikiversity.org for languages that have not reached critical mass of active users. LREforSchools uses the same principles as LeMill for dealing with languages: the user interface is translated into several languages and the content is syndicated from several languages. LREforSchools has a special tag for “Travel well”-resources that do not rely on language.

Encouraging creation of small resources is a goal that is related to reuse and remixing. For finding images and pieces to use in teaching, teachers can rely on Flickr and similar services, but for the purpose of composing a new resource to be shared with other teachers, the parts should have a license that permits that. Hosting suitable media pieces inside the repository facilitates remixing and allows automated attribution. While Wikiversity uses resources from Wikimedia Commons quite naturally, other repositories expect authors to have prepared material that is clear from copyright issues.

5.1 Authoring Tools
In addition to OER repositories, there are authoring tools for learning resources that have repository-like features. Their focus is on creating resources in a certain presentation format, which then can be downloaded or played on site.

LAMS is a tool for authoring IMS LD-compliant learning sequences and has a community where sequences are published and shared (http://www.lamscommunity.org). Playing LAMS sequences requires a dedicated server. Building a sequence is done with a drag-and-drop editor, but as the task is to model a learning scenario, the entry
barrier is quite high. Furthermore, viewing a learning sequence requires the creation of a demo account. The RELOAD editor allows offline editing of IMS LD and SCORM sequences, but it is aimed for more technical users. While IMS LD has promise, modeling learning sequences is beyond LeMill’s goals of supporting teachers in their first steps at collaborative authoring.

5.2 Other Approaches

ALOCOM is an innovative approach for reusing learning objects. ALOCOM is a model about content of learning objects, but instead of providing an online service for composing LOs from LOs, it is used by plugins within popular office software [36]. These plugins try to search the ALOCOM repository for smaller “Content Fragments” that could be useful for the LO that the teacher is trying to create. The LO can then be sent to the repository where it will get automatically parsed into Content Fragments usable by other teachers. By using office software that the teachers already know well, it bypasses many usability issues and in a way provides scaffolding. This approach is highly dependent on automated parsing of LOs and would be less effective for the very heterogenous needs of European teachers. However, doing ALOCOM ontology parsing in LeMill, uploading LeMill content to the ALOCOM repository and fetching Content Fragments from ALOCOM could be a future option.

Also of note is eduCommons (http://educommons.com), a content management system for OpenCourseWare projects. The idea is that a school can have an eduCommons server and host its courses there. These hosted courses are syndicated to other OpenCourseWare servers and repositories. This is a great approach for institutions that can commit to the OpenCourseWare Consortium, as each eduCommons site can set up its workflows and practices as they see fit. But as the consortium targets only higher education, modifying an eduCommons site to serve primary or secondary school teachers would be a design project of its own. In one phase of the project, LeMill was perceived to provide a similar service for basic education: LeMill installations were supposed to form a network with syndicated searches. However, as a lively initial community was a necessity, we concentrated our efforts to one LeMill site.

6 Conclusions

The design of LeMill aims to combine many aspects that are important in authoring and sharing of open educational resources. We have tried to solve several fundamental challenges that hinder European teachers from creating and sharing open educational resources. The information architecture presented in LeMill can be part of the solution helping European teachers share more their learning materials and improve them in a collaborative way.

Thorough validation and evaluation work has shown that LeMill is at least a partial success. From the very beginning of the project we aimed to create a living community—or actually make it a platform for communities. Today LeMill has more than 7,500 members from 56 countries. Teachers from some countries, such as Georgia or Estonia, have their own strong communities, while teachers from some countries like Hungary or Finland have not yet formed a self-sustaining community. Together with Connexions, Wikiversity, and WikiEducator LeMill is currently one of the largest OER initiatives based on peer-produced content.

LeMill has currently over 8,500 learning resources. In November 2007 Ochoa and Duval [37] calculated that average growth for Connexions to be 1.8, for MIT OCW 1.0, and for MERLOT 4.6 new resources per day. In the same time period LeMill’s growth was 3.4 resources per day and in the fall of 2009 the growth rate was 8.0 (not including media pieces).

LeMill has been evaluated by pedagogical researchers working in close cooperation with teachers [38]. The evaluation included authentic trials in schools, national focus group discussions in six European countries, pre-pilots, workshops, and evaluation exercises with teachers between 2005 and 2008. The results were compared to the general ICT capabilities of the teachers with the SIPTEC framework. During the evaluation, one hypothesis concerned LeMill successfully supporting computer supported collaborative learning in knowledge building communities and trialogical learning. This hypothesis was supported by the findings of the study, mentioning the emergence of national communities, peer learning among teachers, extra-cognitive mechanisms, social presence, and the importance of coevolutionary methods in the development of LeMill. Another hypothesis concerning long term impact of LeMill showed promise, but could not be verified in the study as more time was needed to truly see the long term effects [38].

However, the main question for this paper is what part the design of LeMill has played in achieving these results. The design solutions have been

1. making social activities like building collections, building portfolios, discussing and forming groups centered on resources, and having reuse and remixing an explicit part of resource creation,
2. publishing all resources under a license that permits remixing,
3. making metadata creation implicit and manual entry of metadata minimal,
4. mixing languages together and personalizing views based on language preferences, and
5. encouraging the creation of small resources and keeping them open to the larger Web.

In discussions with teachers from several countries one of the main reasons why LeMill has been taken into use by the teachers is its ease of use. It is said to be simple and elegant, while at the same time showing respect for freedom of expression, freedom of assembly, and the existence of different languages and cultures.

However, by looking at the actual amount of collaborative editing in LeMill, the numbers are low. Only 5.5 percent (270 out of 4,890) of resources (excluding media pieces and resources that are clearly stubs, have no tags or language information) have been edited by more than one author. For those resources that are assigned to groups, 9.7 percent (128 of 1,326) have been edited by more than one author. It is a positive finding that assigning resources to groups has a noticeable effect. There are some known factors influencing these numbers. One is that collaborative editing is defined quite strictly: only changes in the actual body text of a
resource count; adding tags or fixing missing metadata doesn’t count. Another is that many of the resources are created in teacher training sessions and while we now try to delete them afterwards, there are still large amounts of low quality content that is generally not visible, but affects these statistics. Our design solutions for supporting collaboration have not created a major change in teachers’ behavior in this aspect.

We may also have been mistaken in our assumption that the smaller granularity of resources lowers the threshold of resource creation. Ochoa and Duval [37] found that repositories that feature full courses have a more active user base than repositories that concentrate on resources of smaller granularity. A course is a natural context for teachers to return and work on. Instead we have had to rely on workshops, competitions, and teacher training sessions to provide temporary meaningful contexts for resource creation.

We have found that it is difficult to maintain the separation between design work in LeMill and “community gardening.” In social software, a design solution can create affordances, but community conventions and introduced rules can override and replace the designers’ intentions. Design can be used to resist certain habits, like the teachers’ preference of keeping unfinished resources private. A designer as a community moderator can do the same thing with a simple discussion post. In LeMill dissemination, one of the designers has held the majority of teacher workshops and gained many important insights into the design problems from teachers. This learning has been mutual: teachers in the workshops have learned about LeMill’s design intentions and in turn adapted to them. This is typical for a “wicked problem”—even when there seems to be progress toward solving it, it is difficult to point out which of the multiple solution attempts actually worked. In the spirit of openness, we have here presented our attempts.

ACKNOWLEDGMENTS

This research was partly funded by ESF grant 7663 and MER targeted research 0130159s08.

REFERENCES


Teemu Leinonen is a designer and professor of new media design and learning in the Media Lab Helsinki at the Aalto University School of Art and Design.

Jukka Purma is a researcher and a software developer in the Media Lab Helsinki at the Aalto University School of Art and Design.

Hans Põldoja is a designer and a research associate in the Institute of Informatics at the Tallinn University and a doctoral candidate in the Media Lab Helsinki at the Aalto University School of Art and Design.

Tarmo Toikkanen is a psychologist, researcher, and software developer in the Media Lab Helsinki at the Aalto University School of Art and Design and is working toward the PhD degree at the University of Helsinki.