Issue 7
2010

# Distributed Repositories for Educational Content

## Part 2: Technology

Michael Klebl
klebl@edu-sharing.net

Bernd J. Krämer
kraemer@edu-sharing.net

Annett Zobel
zobel@edu-sharing.net

Matthias Hupfer
hupfer@edu-sharing.net

Christian Lukaschik
lukaschik@edu-sharing.net

edu-sharing.net e.V.
Postfach 2163
99402 Weimar, Germany

urn:nbn:de:0009-5-27748

**Abstract**

In Part 1 of this article we discussed the need for information quality and the systematic management of learning materials and learning arrangements. Digital repositories, often called Learning Object Repositories (LOR), were introduced as a promising answer to this challenge. We also derived technological and pedagogical requirements for LORs from a concretization of information quality criteria for e-learning technology. This second part presents technical solutions that particularly address the demands of open education movements, which aspire to a global reuse and sharing culture. From this viewpoint, we develop core requirements for scalable network architectures for educational content management. We then present edu-sharing, an advanced example of a network of homogeneous repositories for learning resources, and discuss related technology. We conclude with an outlook in terms of emerging developments towards open and networked system architectures in e-learning.

**Keywords:** Educational content, learning object repository, content sharing, repository networks, web portal

## 1 Introduction

In the first part of this article we reviewed the concept "information quality" (KRCMAR 2005, p. 111) in an e-learning context, where e-learning refers to the use of information and communication technologies in education. Our focus was on digital learning content, including teaching texts, interactive educational software, animations, simulations, exercises and tests. In Section 1 we mapped information quality attributes (such as consistency, avoidance of redundancy, access control and relevance to the target audience) to pedagogical principles, and described how Learning Object Repositories (LORs) provide useful support functions to realize these principles in practice. In Section 2 we discussed typical repository functions and properties and in Section 3 we classified LORs according to a few distinguishing criteria, including content storage model, coverage of topic areas, user and author communities addressed, quality assurance approach, and system architecture.

A review of current practice showed that the majority of virtual learning environments and repositories adopt a centralized architecture. However, different needs and side conditions may forbid a centralized content repository. For instance, institutions or enterprises may want to:

- Integrate the repository with their own IT infrastructure, e.g., to handle authentication and authorization in-house or to embed the repository in their own portal and furnish it with a corporate design;
- Keep the autonomy of operation and responsibility for their own or acquired content;
- Implement their own variant of metadata or establish their own workflows;
- Manage also sensitive documents in their repository, which acts like an enterprise content management system, is accessible in the intranet, and needs to function independently from internet connections.

The open education and open educational resources movements (see Part 1 , Section 1.2) have a global demand that cannot be addressed by a single common repository. These movements raise the necessity of federated content management solutions because different content providers prefer different repositories. Similarly, in commercial environments the networking of repositories is a key requirement. For example, if you offer educational content to a company, you may also want to integrate commercially purchased learning materials that are drawn from education providers' and publishing houses' own repositories.

Federated and distributed repositories differ in various technical and usability aspects as will be explained below.

## 2 Federated Repositories

The number of repositories for e-learning applications is increasing, both in academic and commercial contexts. Though early implementations were independent, isolated content silos have limitations that can be overcome by networking the different products. One type of networking, federated search, enables search processes beyond the confines of a single repository.

## 2.1 Federated Search

Searching in networked repositories is called federated search and the network of heterogeneous repositories is called a federation. The repositories participating in a federation are autonomous. They run on different, usually geographically distributed servers, and can communicate with each other via the Internet or some intranet. Federated search engines can search content that is not available to web search engines. The repositories participating in a federation are autonomous. They run on different, usually geographically distributed servers, and can communicate with each other via the Internet or some intranet.

A federated search requires agreement on a common set of metadata and a common protocol for search queries. Examples include: Search/Retrieve via URL (SRU) for libraries (LIBRARY OF CONGRESS, 2008) and XML Query Language (XQuery) for XML databases (W3C, 2006). To ensure the interoperability of federated searches in repositories, the Open Archives Initiative (OAI, 2002) has standardized the “Protocol for Metadata Harvesting” (OAI-PMH).

Searching in federated learning object repositories was first promoted by the Ariadne Foundation, developers and operators of the ARIADNE Knowledge Pool System (DUVAL ET AL. 2001), the MERLOT (1997) project, and other groups. In 2005, this consortium of LOR operators presented a solution based on a common query interface, SQI (Simple Query Interface), and a three-layer software architecture (SIMON ET AL. 2005). The bottom layer includes the actual repositories and performs persistent data storage functions. The middleware layer comprises a search engine and a registry in which information about the connected repositories is kept. The application layer provides search functions that are embedded in portals or are available via plug-ins in the learning management systems. The search engine distributes incoming queries to the participating repositories, gathers the hits, sorts them, and converts them into a standard format. Via the application layer, the compiled result is then presented on the inquiring client system.

This initiative has triggered the formation of a worldwide alliance, GLOBE (2008), whose aim is to disseminate digital learning resources to as many teachers and students as possible. One of the means to achieve this goal is the federation of many repositories.

## 2.2 Benefits and disadvantages of federated learning object repositories

Although searching in a federation of repositories delivers a standardized presentation of all hits, the heterogeneity of participating sites becomes obvious when accessing remote content containers. They have different user interfaces, organizational structures, and export functions and formats. In addition, the usage rights and rights administration functions are rarely harmonized. In a nutshell: heterogeneity makes life difficult for users.

Let us take a look at an example to illustrate the heterogeneity issue. A federated search for the term “web service“ from within the MERLOT portal delivers about 10 pages of different results. These include:

- Online teaching material entitled “The Web Services Value Chain”,

- A simple website with a link to a web service to validate HTML pages,
- A Mathematica object, which can be downloaded and played locally and which assumes that the user's PC has a Mathematica player installed,
- A link to a slide presentation distributed via Slideshare, and others.

For all presumably useful hits the users must follow the links and, if the content is promising, grapple with the format of the resource and the usage rights according to the special features of the respective repository or referatory.

Federated repositories allow only a restricted form of user administration. In the simplest case, all content is accessible or at least all registered users in a repository federation have access to the content in the network. A more detailed approach is impossible because of the different solutions in handling user rights.

The disadvantages of a federation as opposed to a central repository are offset by a crucial advantage: in the past, institutional users of educational content, such as universities, publishing houses, and companies, have shown no willingness to make their content available in a central, third party-operated repository. In a federation, they can operate their repository autonomously and still cooperate with others. These benefits and other basic features of distributed repositories will be described in more detail in the next section.

## 3 Distributed Repositories

Distributed repositories share the benefits of decentralization and operator autonomy with federated repositories, but avoid the disadvantages resulting from the heterogeneity of federated systems. Compared to centralized repositories, distributed repositories can adapt better to growing numbers of users and increased content. However, they are more difficult to implement because they present new design challenges, as described in the remainder of this section.

### 3.1 Architecture and functionality

A distributed repository is an amalgamation of independent installations of the same repository. Users see a distributed repository as a single system although the data is physically distributed among several repositories. Generally, each site operates an installation of the repository software autonomously on a separate server. At each site, the respective repository can be embedded in other authoring systems and learning environments via standard interfaces. Each site repository is linked with the other repositories of the distributed architecture via specified interfaces and must comply with the agreed standards for user authentication and authorization.

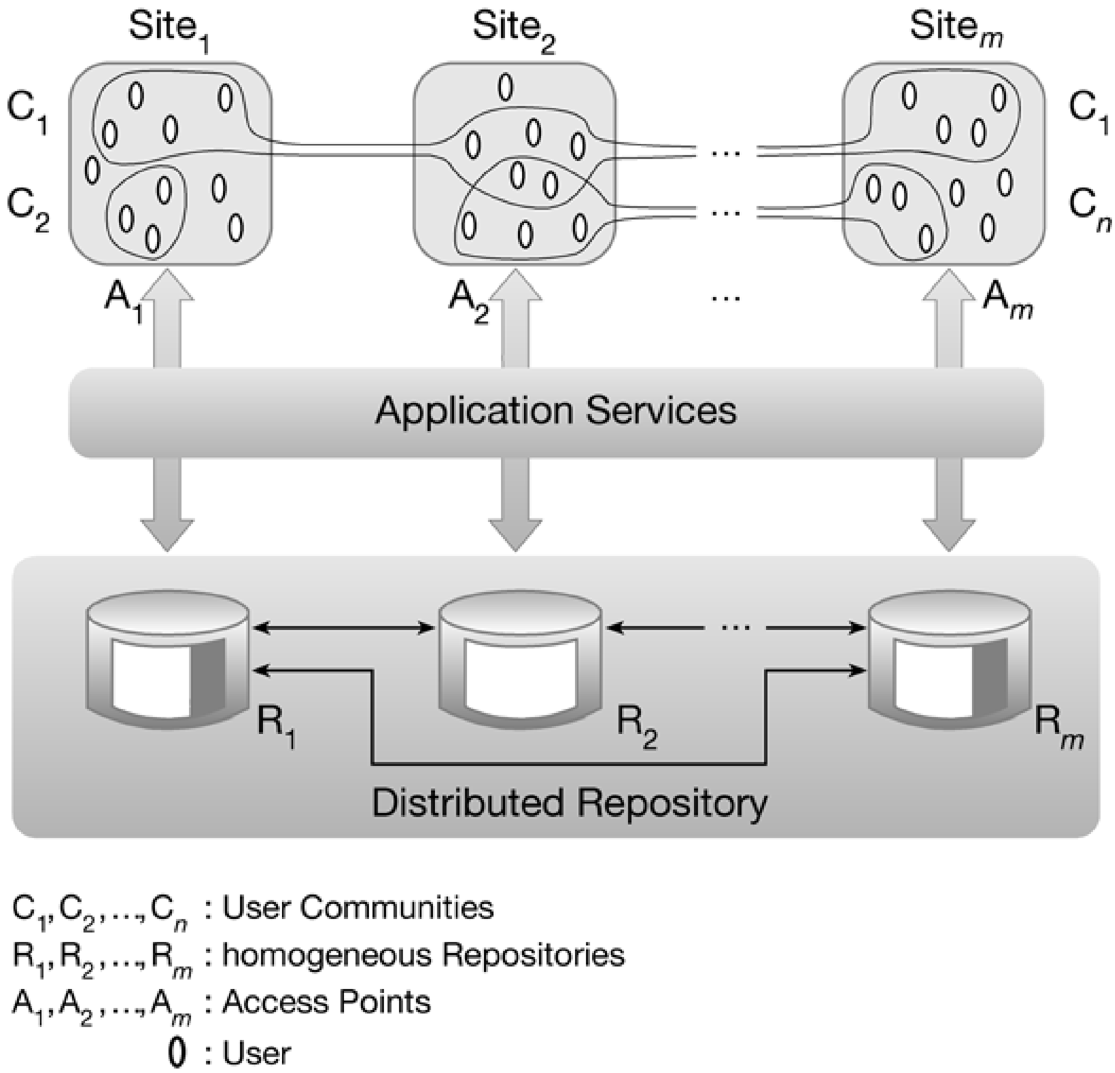



*Figure 1: Distributed repository with cross-organizational communities*

The distributed repository in Fig. 1 has m sites and m access points through which users can access the complete system. All sites offer the same application services for LOR access (see also Fig. 1 in Part 1). Behind the access points are homogeneous LORs that are linked with each other. The autonomy in distributed repositories allows the participating organizations to manage open content (depicted as white areas) and closed content (gray areas) uniformly. The operating institution also decides autonomously which closed resources are made available to local and remote users. Autonomy also creates flexible options for cross-institutional cooperation and dynamic group formation. User communities (C1, ..., Cn in Fig. 1) that want to share specialist interests and cooperate with each other can be formed across access points.

A distributed LOR scales better than a centralized LOR. The latter runs on a single server and is operated by a single facility, while the performance of a distributed LOR can be adapted to increasing demands by adding more servers.

## 3.2 Transparency

A distributed repository looks like a single system to the outside. This property is based on different attributes of the (IT) concept of transparency, which means that certain properties of entities remain hidden.

**Location transparency** requires that users access an object via a link, regardless of which specific repository is managing the destination object. Access transparency is when a user always accesses objects in the same manner, regardless whether they are stored locally or remotely. A federation of repositories shares the location transparency property but does not exhibit access transparency.

**Replication transparency** allows data (or portions thereof) to be copied many times on different sites, while to users it appears there is only one copy. In Section 1 of Part I we argued that redundancy impairs information quality because it induces inconsistencies when redundancy occurs unorganized. For distributed systems, redundant data storage is useful to ensure availability of the resources if individual servers fail (failure transparency), or if the access speed needs to be increased by load balancing. Distributed systems management cares for data consistency by updating different copies of the same data periodically. It is worth noticing here that only near consistency can be guaranteed in a federated or distributed system because no global control exists and in between two updates local changes can be made. Another example of replication is maintaining the metadata and the index of the complete system on all peers to optimize the performance of search processes. This index is updated regularly in the background to include local changes in the complete system. Such redundancies, coupled with concurrent changes to data in distributed repositories, prevent complete and continuous consistency. In general, the synchronization measures can create only near consistency.

With distributed repositories the components of a collection of resources can be managed in different repositories. The property of fragmentation transparency ensures that the physical distribution of a collection of content is hidden from the users of this collection and appears like a local collection.

**Concurrency transparency** ensures that several users who access a repository at the same time do not affect each other. Without suitable protective measures, two users of a site could generate revisions to an object independently and, upon saving their changes, overwrite the other's revision or even produce a corrupted revision from a mixture of both changes. Such effects can be prevented on a single site with lock mechanisms, such as checkout and check-in. However, concurrency transparency is not desirable for all types of uses. For example, if the LOR also offers functions for synchronous group work, such as cooperative editing of resources, the users working concurrently on the same resource must know about each other. To enable a high degree of parallel work, more elegant methods are needed than the above-mentioned lock mechanisms.

Lock mechanisms to implement concurrency transparency apply only locally. If different copies of an object exist in different repositories of the network, multiple users could work independently of each other on revisions of different copies of the same object. When revisions are imported, they can cause **conflicts between the revisions** and thus inconsistencies in the complete system. To prevent such inconsistencies, mechanisms must be installed to resolve conflicts. A rather inelegant conflict resolution strategy is to prompt the authors of concurrent revisions and let them decide which revision will survive. Another strategy might rely on priorities associated with authors and give right to the change of the author with the highest change priority.

For users of LORs it is important that they can assess sources of information before using or buying it and also present the corresponding resources in their learning management system. For web-supported applications and LORs, we call this specific transparency property **rendering transparency**. This requires the system to provide a rendering service that can present objects of every resource format managed in the LOR on the intended presentation media (screen, audio output, DVD).

## 3.3 Distributed user administration

In a distributed repository, the tasks of creating new users, changing user profiles, and deleting registered users is more difficult than with a centralized solution. The following options exist:

- **Manual multiple registrations**. In the simplest case, users have to register separately on each installation. This can have undesired side effects, such as inconsistent profiles and roles for one person on different installations. In addition, users have to authenticate themselves several times when accessing different sites, which counteracts transparency.
- **Central users register** and **trusted communication**. One way of avoiding the disadvantages of multiple registrations is to keep a central users register through which all accesses are filtered. There must then be a trusted communication connection between the server that keeps the register and the servers that provide the content. This connection is used to transfer the access request from the register together with the user ID and role and rights information. This information can, e.g., be communicated through a certificate. The downfall of this approach is a centralized system with content balanced on a set of servers.
- **Transparent multiple registrations**. Alternatively a registration can be tagged as “original” and the associated information can be distributed to all sites as soon as the user has been set up and when any changes are made to the profile. This alternative also needs a trusted communication connection but avoids the bottleneck of a central register. Instead of distributing an original registration automatically to all repositories, a lazy process can be used where the registration is propagated to other repositories only when they are first accessed, transparently.
- **Federated identity provision** and **single sign-on**. The open source software SHIBBOLETH® (2010) is a prominent solution for a single sign-on service within and across organizational boundaries. The solution relies on a federation of identity management systems, the users’ home institutions that manage their own authorization data and policies. Upon a user’s access request to an online resource, two situations are possible. If the user has recently accessed another protected resource and thus operates in a single sign-on session, access is granted immediately if this user owns the proper access rights. Otherwise, the user’s web browser gets redirected to the single sign-on service, which prompts the login page of her home

institution. If the authentication at the home site succeeds, the single sign-on service passes relevant user attributes to the resource server, which decides whether the access is granted. As long as the session exists, no further authentication is required.

The common goal of the last three solutions is to avoidnexplicit multiple authentication steps. To keep the complete system (quasi) consistent when changes are made, some administrative effort is needed in the background.

## 4 Extensive Networking of Learning Object Repositories

In the previous sections, we discussed the amalgamation of heterogeneous repositories into a federated system and the amalgamation of geographically dispersed homogeneous repositories to a distributed system. A further step is a hybrid architecture that combines both approaches.

To outweigh the deficits of federated systems – accessibility and access control – we suggest a combination of federated and distributed architectures in such a way that the access points of the distributed system take over critical functions such as searching, browsing, selection, rendering, and trusted user and group administration. Other functions of repositories like long-term, secure storage of media objects can still be handled by heterogeneous repository components. In other words, the core of a combined architecture is a network of homogeneous repositories. Third party repositories, which are typically of heterogeneous type, are connected to this homogeneous network via at least one access point. The content of the federated repositories remains where it is, the homogeneous network indexes it for an effective search and an integrated selection. The homogeneous network then also allows a preview and controlled access that comply with the requirements of location transparency, access transparency, replication transparency, and fragmentation transparency.

### 4.1 Using learning objects in combined architectures

Rendering transparency is also important for a combined network of distributed and federated systems in terms of the different usage functions like provision, preview, and integration into virtual learning environments. It should be possible to at least preview learning material and learning arrangements found in the connected repositories. In the best-case scenario it should also be possible to integrate these directly into the learning processes, e.g., through a virtual learning environment. As mentioned in Section 3.2, this would need suitable rendering services for every format stored in the resource. The access points in the network must not only manage content, but must also offer rendering services. For a resource format that cannot be directly presented in the web browser, they need to find and execute a suitable rendering service. These rendering services can also be distributed among the connected systems. For a proprietary format, such as is the case with the Connexions repository (BARANIUK 2008), the rendering service will probably be executed within the scope of the Connexions repository. But for standard formats, like SCORM or IMS QTI, the access nodes of the homogeneous network will provide rendering services directly.

## 4.2 Presenting learning content: flexible rendering services

New or updated rendering services should be easy to register and manage. This enables new presentation formats to be integrated easily, such as special formats for mobile devices or formats that suit the graphic design guidelines of an organization. Flexible management of rendering services also offers advantages for load balancing and specialization. Distributed systems can handle high access volumes, while other servers are specialized for specific requirements (e.g., for streaming time-based media). Rendering services also open up access to learning material and learning scenarios outside the closed course environments of established virtual learning environments. They can also be integrated into and used in personal study portals (in terms of personal learning environments, PLEs), as long as the corresponding access rights exist. Tracking services are also needed for use in real teaching and learning processes. They gather the usage data (e.g., test results) and report them to the learning management system (LMS).

## 4.3 Rendering services and access administration

For combined architectures that can combine freely accessible and licensed learning materials, recipient-oriented management of rendering services is important for differentiated access administration. For example, the rendering service can be limited to a preview that forbids downloading and integration into virtual learning environments. Specific formats like SCORM or IMS QTI (or also complete courses in the format of a virtual learning environment) can be unpacked and presented after the search, but they cannot be used for other purposes. Standard formats like PDF and office documents or image, audio, or video files can be presented for a preview only in excerpts or in a lower quality to allow access to the high quality media objects only if it is granted.

## 4.4 User administration in combined architectures

While standalone repositories enable access to non-public content by storing a password for the repository in the virtual learning environment, advanced networking of repositories should provide a connection based on the principle of mutually trusting systems. Simpler solutions attempt to map the rights of the virtual learning environments to the repository. But this complicates the connection of several different types of virtual learning environments since they usually have distinct rights and role concepts. More flexible are solutions that establish internally a separate rights system for the authors and content managers and define the access authorization from third party systems within the scope of interface implementation. In this way, every virtual learning environment or every other connected system (e.g., study portal, library system) is given permission according to the respective role and rights concept. A user of the repository (usually a teacher) inserts content in a learning that he or she can access. When a learner accesses the content, the repository checks whether the learner is actually a user of the virtual learning environment and is logged into the course for which the content was released by means of a double handshake. This principle of trusted systems is employed in the edu-sharing repository network (see Section 5).

## 5 edu-sharing: Distributed LOR and Community of Practice

The methods and base technologies of the repository network edu-sharing were created by the project CampusContent (CAMPUSCONTENT 2010). The German Science Foundation (DFG 2010) funded this project as a Competence Centre for e-Learning between March 2005 and July 2009. As a network of homogeneous repositories, edu-sharing provides a technical infrastructure for managing, networking and marketing educational content and codified know-how. It also allows the integration of content and pedagogical know-how into learning and educational workflows. In the following subsections we briefly review the history of the project, sketch the core functionality of edu-sharing with reference to the features discussed in Section 3 and 4, and present the organization of user and developer groups through the non-profit association edu-sharing.net (edu-sharing, 2010).

### 5.1 From CampusContent to edu-sharing

The repository network edu-sharing is an outcome of several years of R&D effort. The CampusContent team at FernUniversität in Hagen performed this research. The project CampusContent aimed at the design and construction of solutions that would address two challenges:

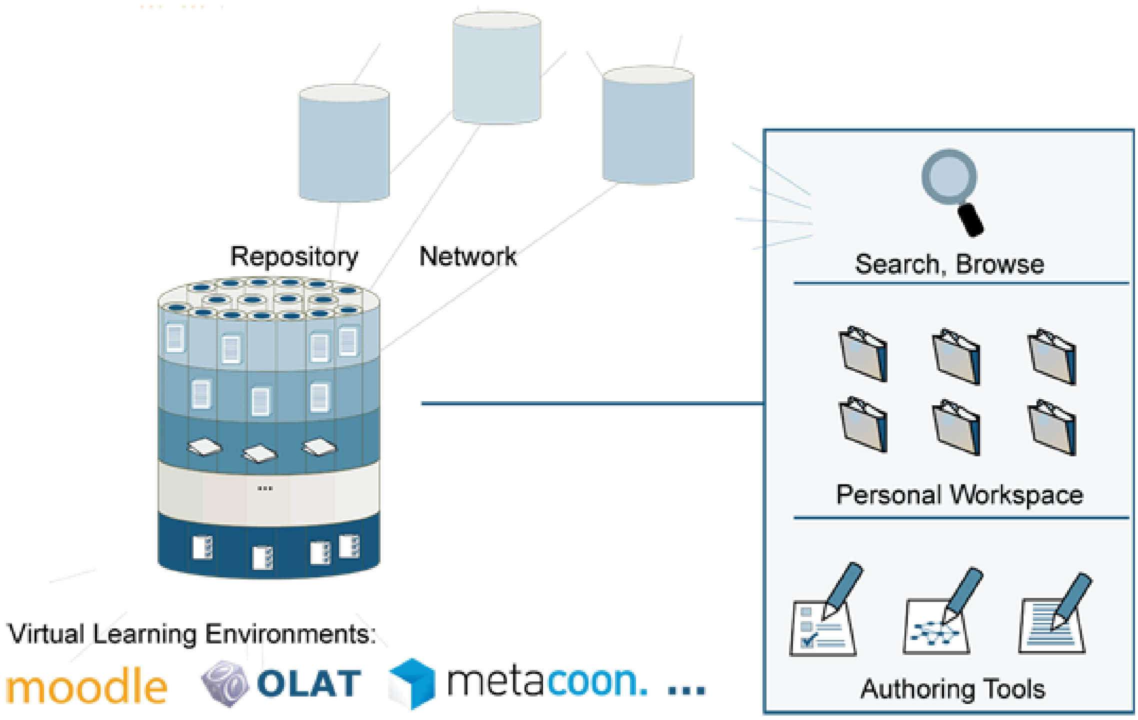
- Secure the usability and sustainability of digital learning materials and mature teaching methods developed in daily practice.
- Provide a systematic and easy-to-use solution that allows content users to integrate resources from different providers and reuse them in different pedagogical settings meaningfully.

Originally the CampusContent team focused on the development of added value for universities and science. Empirical research including user surveys and self-assessment revealed, however, that development and sharing of large knowledge- and content-bases requires a widening of focus to address any form of education. In addition, the team became aware of a high demand for a content sharing infrastructure that bridges heterogeneous learning management platforms from schools, vocational colleges, and continuing education institutions. To cope with this widening of target groups, the name edu-sharing was chosen both for the association and the repository network.

### 5.2 Survey of edu-sharing

**Cross-institutional search** and **resource sharing**. edu-sharing is a distributed LOR in the sense of Section 3.1. It consists of a set of interconnected sites as depicted in Fig. 1. Each site maintains its resources in its own instance of the edu-sharing repository, which is embedded in a set of locally preferred authoring and learning tools (see Animation 1) and may be connected to other edu-sharing repositories. Thus the repository network allows cross-institutional search and exchange of learning material and pedagogical scenarios.

**Transparency**. As users decide which of the connected repositories should be queried in a search, edu-sharing does not offer location transparency. Access and rendering transparency are supported, however. Currently no replication strategy is applied because the supply of resources and the number of users are still limited. For the same reason, no decision has been made about a conflict resolution strategy.




*Animation 1: edu-sharing repository with attached end-user tools*

*(when clicking on the image, you will be led to a website illustrating the mode of operation of edu-sharing's search function, workspace and selected authoring tools included in the standard distribution of the edu-sharing open source software. On YouTube you can find further demonstrations with German explanations about edu-sharing's search functionality, the interplay of edu-sharing and Moodle, personal workspace, and an integrated editor for creating and adapting learning scenarios.)*

**User management**. Users are managed through federated identity provision and single sign-on. This way, universities, schools and commercial education providers continue to manage their own users autonomously. An access to a protected resource maintained in (for instance) closed publisher resource pools can be routed through edu-sharing exactly the same way as an access to another edu-sharing site (see Fig. 2).

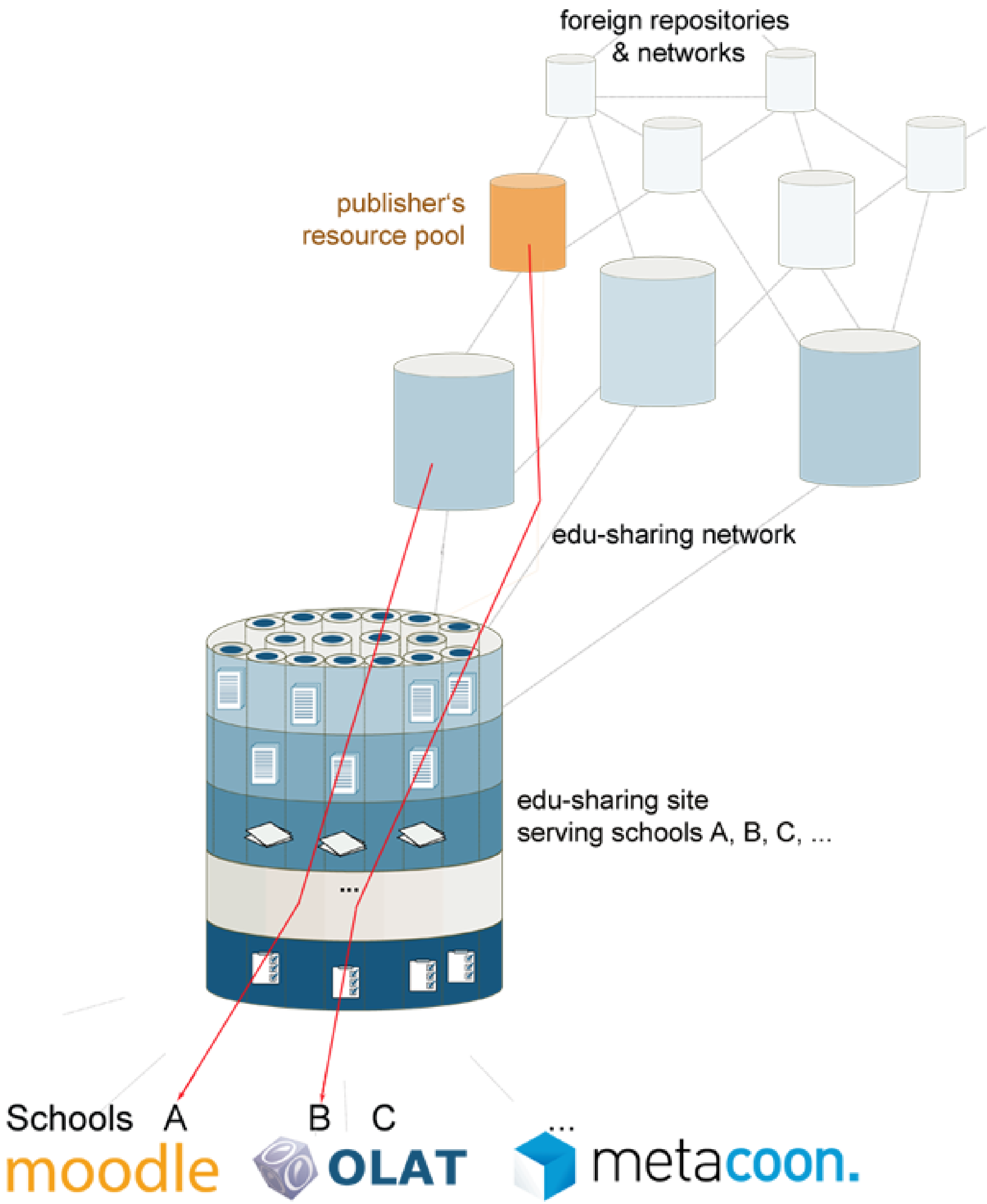



*Figure 2: Accessing online resources within and beyond an edu-sharing network*

**Personal workspace**: Each registered edu-sharing node owns a personal workspace (Animation 1). Using a combination of file manager and mind map (associogram), the workspace forms the innovative graphical interface of a personalized document management system. Here published and unpublished objects are stored, arranged in folders and shared with or worked on by invited users.

**Metadata**. The repository network stores optional file types and links on external resources together with their metadata. The available metadata sets (Dublin Core and LOM) can be amended to complement the organizational or educational area. They can also be modified. To support the special requests of German schools, a modified LOM-set, LOM DE, has been added to sites managed by communal computing centers.

**Data model**. The edu-sharing repository network manages learning objects with varying granularity. It handles complete courses and reusable learning paths that model these learning arrangements (KRÄMER ET AL. 2010). Figure 3 shows the data model underlying edu-sharing repositories. In simpler form – without the scenario part – this data model is used by all LORs. A distinction can also be made between the elementary content and collections of such content; for example, individual media objects in the LOR as opposed to a course containing media objects from the repository. Both are associated with metadata that optimize finding and evaluation. Some repositories also maintain relationships between elementary content, which is generically modeled by the relationship “is related to” in Fig. 3. For instance, if versioning is supported – as it is in the edu-sharing repository network – this relationship specializes into “is a revision of”. For edu-sharing specifically, the differentiation of the term “content” into objects and pedagogical scenarios is depicted at the bottom of Fig. 3. Recursively, a scenario can be made up of smaller scenarios and scenarios can refer to objects from the repository.

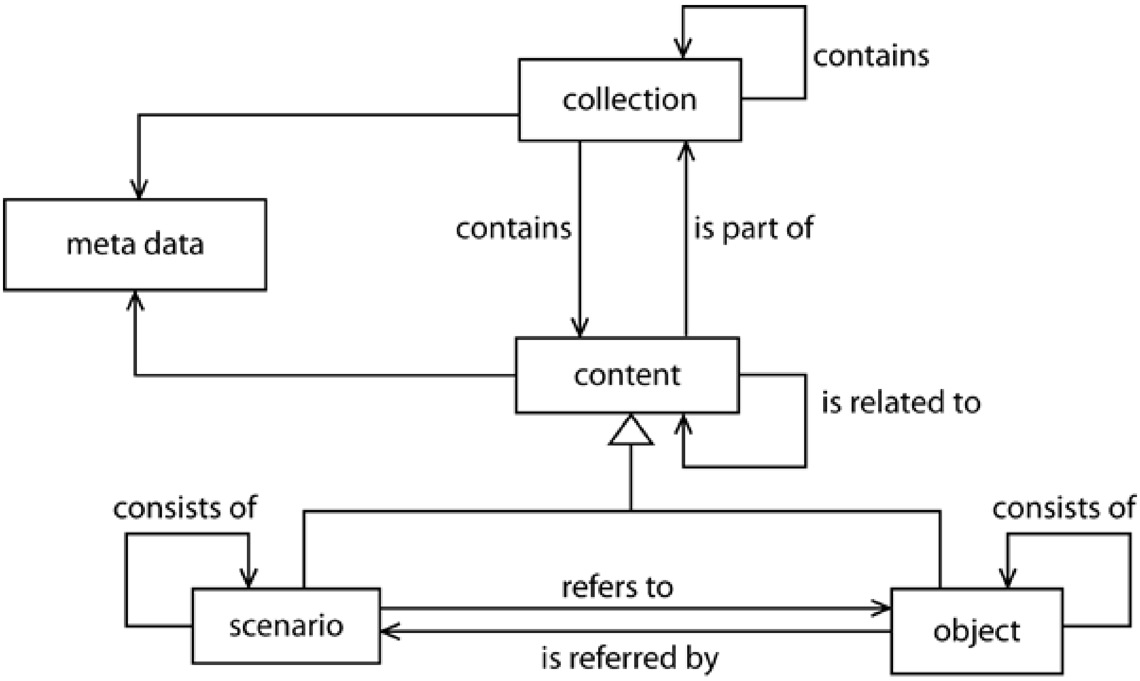



*Figure 3: Data model of the edu-sharing repository network*

**Integration of virtual learning environments**. edu-sharing repositories can be linked with existing LMS technology. Currently, interfaces with the widespread learning environments Moodle, metacoon and OLAT exist. The integration of Fronter in a simple form is planned for early 2011. Different integration levels have been discussed with the Fronter team:

1. use edu-sharing content within your LMS using edu-sharing’s web services;

2. upload material (pdf, scorm, media, …) into edu-sharing instead of an LMS folder and reference it from there; different alternatives are possible to achieve this level:
a) add a central folder for edu-sharing content to the LMS's central material folder;
b) add an edu-sharing home folder to the LMS's home folder;
c) add edu-sharing course or group folders to the LMS's course or group folders, resp.;
3. upload courses, add a rendering plugin to make your own LMS course format viewable for other LMSes;
4. enhance your rendering plugin such that users of other LMSes can view your own LMS courses.

**Standard distribution**. Besides the integration of popular virtual learning environments, the open source software edu-sharing comes with a collection of authoring tools, including a QTI 2.1 compatible off-line editor for exercises and tests, a SCORM-compatible offline editor for courses and course fragments, and an editor for pedagogical scenarios (KRÄMER ET AL. 2010). Through open interfaces, an institution's preferred authoring tools can be integrated to replace these default tools.

**Installation and adaptation effort**. The installation and configuration of the core components of edu-sharing takes 3-4 person days for qualified IT personnel. The integration of an existing LMS for which a specific integration interface already exists (Moodle, metacoon, OLAT) may take another half day. If no such prior integration experience exists, 4-8 additional person days are required, depending on the level of integration desired. The effort for maintenance and upgrade management compares with similar server-based applications. The migration and adaptation of legacy resources (metadata and content objects) requires several person weeks or even months, depending on the volume of data, their consistency and their degree of agreement with the new data model. The rollout of an edu-sharing site is governed by user and documents management regulations that adhere to local needs and the terms and conditions of the edu-sharing network community.

**Benefits of edu-sharing**. Using edu-sharing as a central storage medium rather than saving content in an LMS directly copes with different drawbacks of conventional solutions: a) an institution operates more than one LMS but its users want to share resources; b) an LMS is unable to manage resources across different courses, which requires the handling of multiple copies; c) other systems like authoring tools or project libraries need to access content. If edu-sharing is used as a central digital content management system in an organization, it is likely that existing content pools need to be integrated. Schools in the state of North Rhine Westphalia, for instance, have been using long existing media pools like EDMOND (2010). This and other foreign content pools have been successfully integrated with edu-sharing. The first version of an integrated media search engine for schools that retrieves content from different pools has been realized in the project LEARN:LINE NRW (2010).

**Base components and development tools**. edu-sharing's functionality was designed to build on mature base components and integrate well with existing IT infrastructures. Java was the preferred implementation language. The standard distribution of edu-sharing integrates two Java-based open source (web) content management systems: LIFERAY (2004) serves as a portal in which all edu-sharing functionality is embedded and

ALFRESCO™ (2005) provides the core resource management component. An edu-sharing peer can be deployed either to a central authentication service (for example, LDAP) or to a single-sign-on service (e.g., CAS). Relying on appropriate open source systems has the advantage that the edu-sharing developer community can focus on a usable GUI and e-learning specific requirements, while other developer communities maintain and further develop edu-sharing's base components independently. The edu-sharing user interface was built with GOOGLE WEB TOOLKIT (GWT, 2006). GWT also uses Java as implementation language. It generates HTML and JavaScript, and largely masks browser differences. A GUI with its controls can be parameterized to allow its easy adaptation to different application contexts.

Liferay was chosen because more and more educational organizations use it. It supports the Java Portlet Specifications JSR 168 and JSR 286 and includes useful portlets. With these portlets institutions can combine a wide range of applications for a personalized online desktop for users (teachers, students, administration co-workers). Typical applications include: course management and registration, LMS functionality, further administrative software, mail services, and website management.

Alfresco was selected among several alternative content management systems including Apache Slide, Apache Jackrabbit, Fedora Commons and others. The internal evaluation took place in 2007 and put emphasis on the following criteria:

1. Stability, size of user and developer communities;
2. Functionality, usability and value for the user;
3. Assigned quality of the software (and value for the developer).

In addition, Alfresco has a resilient architecture, supports the Content Repository Specification JSR 170, offers plugins for integrating portals and authoring tools (like office applications), and provides APIs for various programming languages.

In the hindsight, both Liferay and Alfresco apparently have been good choices. In a recent market share report of open source web content management (WCM) systems by WATER & STONE (2010), Liferay leads the Java WCM market and Alfresco is not far behind.

**Content rendering**. A media rendering and conversion service permits the direct replay of repository contents. For this purpose the service converts, where necessary, content into playable formats. Thus foreign or older formats can be made permanently accessible. Currently, the rendering service uses modules for the reproduction of an abundance of graphics, sound and video files, for QTI-compliant tests and exercises, and for SCORM-courses. Moreover, a metacoon or OLAT user can play Moodle courses or course components from within his or her LMS. Similarly, publishers and other content providers can host and play their content for a client's distant LMS systems on their own servers. In this way, an access control is possible, and functions protecting against copying can be integrated into the rendering service.

## 5.3 Open Innovation and Community Building

**Sustainability**. Publicly funded research projects that aim to build software prototypes have recurrent problems with the development and sustainability of their technology once the funding period ends. This is particularly acute for university projects that lack the necessary human and financial resources to bridge the chasm between a research prototype and a robust and usable product. To address this challenge, the strategy of the CampusContent project includes several components: open source software, custom-support for target groups, multiple content licenses, linking with third-party repositories, quality control, community building, and knowledge transfer.

**Open source software**. The benefits of open source software include its low acceptance barriers for adopters of the technology who can use, modify, and redistribute the software without paying fees, and its potential to be accepted by a community and to evolve through community cooperation.

**Support for target groups**. Potential users of edu-sharing include: (1) educational institutions, enterprises and research institutes, who can jointly develop and manage learning materials, knowledge content and research results; (2) authors, who want to gain exposure to a larger audience or who want to co-develop, manage and share teaching and learning materials with their peers; (3) instructors, who want to find and combine repository content and employ it in their teaching; (4) publishers and other content providers, who can link their content bases with edu-sharing to provide access to their licensed assets for admitted edu-sharing users without committing their assets to the public at large. To these groups, edu-sharing offers a repository network that manages published, variably licensed teaching and learning content and codified methodological know-how. Beyond this, it provides a range of authoring tools and sharable, user-configured workspaces in which users can manage and share unpublished information with invited peers.

**Multiple content licenses**. Initially the CampusContent project focused on open content licenses, in particular, CREATIVE COMMONS (2010). Such licenses adhere to similar principles as open source licenses in the sense that content remains free for re-use, can be combined with other content or can be changed under certain conditions. The rationale behind open content is the feedback and quality control loop to authors and re-users that is established by peers who evaluate, reuse and improve content and by users who report on their learning experience. From edu-sharing users we learned that open content licenses do not always apply to attractive content because of copyright restrictions. To enable the sharing and reuse of such resources, the edu-sharing.net corporation is investing in legal advice to establish more restricted license models and thus reduce the effort for individual user communities. The built-in license manager is flexible enough to manage multiple license models.

**Linking with third-party repositories**. It is unlikely that publishers and other institutions hosting educational resources are willing to entrust their assets for storage and publication in the edu-sharing network. To provide access to foreign resources from within the repository network, third-party repositories have been linked with edu-sharing based on trusted server communication protocols and federated identity provision (see Section 3.3). Examples in the school sector include SODIS (2009) and EDMOND (2010). In both cases, metadata were quality assured, filtered and then replicated.

**Quality control**. The edu-sharing community promotes three forms of quality control for content and its organization: social software, quality controlled knowledge maps, and didactic scenario templates. We are currently implementing content rating, recommendation, and annotation functions that will allow individuals and communities of practice to evaluate resources maintained in the repository network, aggregate the ratings of different users and communities, and derive recommendations based on reuse and context information. Ratings can then be used as additional filter in search processes. Ratings and annotations are also available to content authors as feedback information.

Knowledge maps serve as an orientation aid and an efficient mechanism to find resources in certain disciplines or curricula. Such maps are particularly acute in the school sector in which curricula contents are largely predetermined by authorities. Qualified editors are currently developing knowledge maps for selected school subjects and class levels. The Swiss project Educational Landscape Psychology (EDULAP, 2008) aims to develop a technical system that supports the search for educational resources and is based on so-called orientation maps. Cooperation contacts have just been initiated.

Didactic scenario templates are codified learning and teaching scenarios of different granularity (KLEBL ET AL., 2010). They are defined in an abstract form without reference to specific resources and tools to make them applicable to many subject disciplines. Fine-grained examples of such scenarios include devil's advocate, active structuring, flashlight, brainstorming, concept mapping, think-pair-square and web-quest. More complex scenarios, which often rely on tool support, include case study, jigsaw classroom, strategic problem solving, or project-based learning. Experts have defined and explored such scenarios, we have codified and published them as templates in edu-sharing and we have reused selected templates in specific courses (see, e.g., KRÄMER & KLEBL 2009).

**Community building** is a key element of all open source and open content projects. In our context, the term community encompasses the notions of shared concern or passion about a topic (WENGER 1998) and relationships that develop through interaction and activity of distributed individuals and groups who are connected through ICT. To trigger the creation of a community of practice of content creators and re-users, the CampusContent team regularly organized forums to report about objectives and advances in the project, published project results at international conferences and presented prototype versions at fairs. To ease the formation of a developer community, certain parts of the development work were outsourced to software companies. This way, external developers were introduced into the system's architecture and source code. In addition, third-party open source components were integrated, which gives rise to open innovation (CHESBROUGH 2003). After the open source software was publicly announced during LearnTEC in February 2010, a non-profit association, (edu-sharing.net) was founded as an independent means to coordinate the ongoing innovation process. This association has defined how corporations who operate an edu-sharing repository can link to the repository network and set up an organizational framework for user and developer communities to collaborate effectively.

**Knowledge Transfer**. As a community service, edu-sharing includes an information portal that provides information about e-learning topics for consultants, educators and authors. Together with partners, such as the German Initiative for Network Information (DINI), the Association of German University Computing Centers (ZKI), and e-learning experts affiliated with universities and service providers, this information portal is continuously updated. The portal includes the following: recommendations for producing reusable content; information and guidelines about e-learning and authoring tools; situation and problem-oriented consulting information for educators; templates and best-practice examples for learning content and adaptable learning scenarios; hints about legal issues; and much more.

## 6 Related Work

The landscape of learning object repositories and referatories is rich, and cannot all be described in depth. Instead, we focus on representatives for the typology presented in Section 3 of Part 1 and Section 3 of this part.

**Connexions** ( www.cnx.org ) is operated centrally at William Marsh Rice University in Houston, Texas (BARANIUK 2008). It has no thematic limitations and includes the entire education area of schools and universities as well as vocational training. Connexions is a repository (i.e., the digital learning material is kept inside the system itself) with a centralized architecture. Community aspects are the focus of Connexions: teachers and authors can organize themselves into interest groups in the social network. The team behind Connexions promotes the open education movement: all content is open and can be read without prior registration; for write access, registration is necessary. Content-wise it is organized in modules, each covering a specific topic. A module (corresponding to a course) typically consists of a module text and comes with additional media assets. The text is represented in CNXML, an XML-based format common to all module texts. Connexions offers an authoring system producing CNXML output. It is transformed into HTML for rendering in browsers. Media assets are maintained separately in a zip archive and are viewed separately. Module metadata, which also keep information about linkage of modules, are maintained in database for efficient retrieval. Connexions provides so-called lenses in the form of social software that allows reviewers and editorial bodies to identify high quality content.

**MERLOT** is a prominent referatory operated by California State University (Long Beach, USA) (MERLOT 1997). It covers a wide range of academic disciplines. All items maintained undergo a review process by a central editorial board to ensure the quality of the learning material offered. Access to the learning material is open. Content is stored in multiple heterogeneous sites prohibiting location and access transparency.

**Copendia** ( www.copendia.de ) is a web-based marketplace for e-learning offerings. The German Ministry of Economics and Technology (BmWT) sponsored Copendia. It primarily markets content for advanced vocational training. Copendia provides authoring tools and learning management systems for creating and using content. It is thus a centrally organized repository focusing on content from education providers.

The **LON-CAPA** system from Michigan State University (East Lansing, USA) is a distributed system for managing and evaluating content (cf. KORTEMEYER 2009). Individual installations of the LON-CAPA system are linked with each other and allow content to be shared. Each installation includes a repository and a learning management system. LON-CAPA supports closed user groups; content is not accessible for the general public. Within LON-CAPA, individual disciplines form specific user groups.

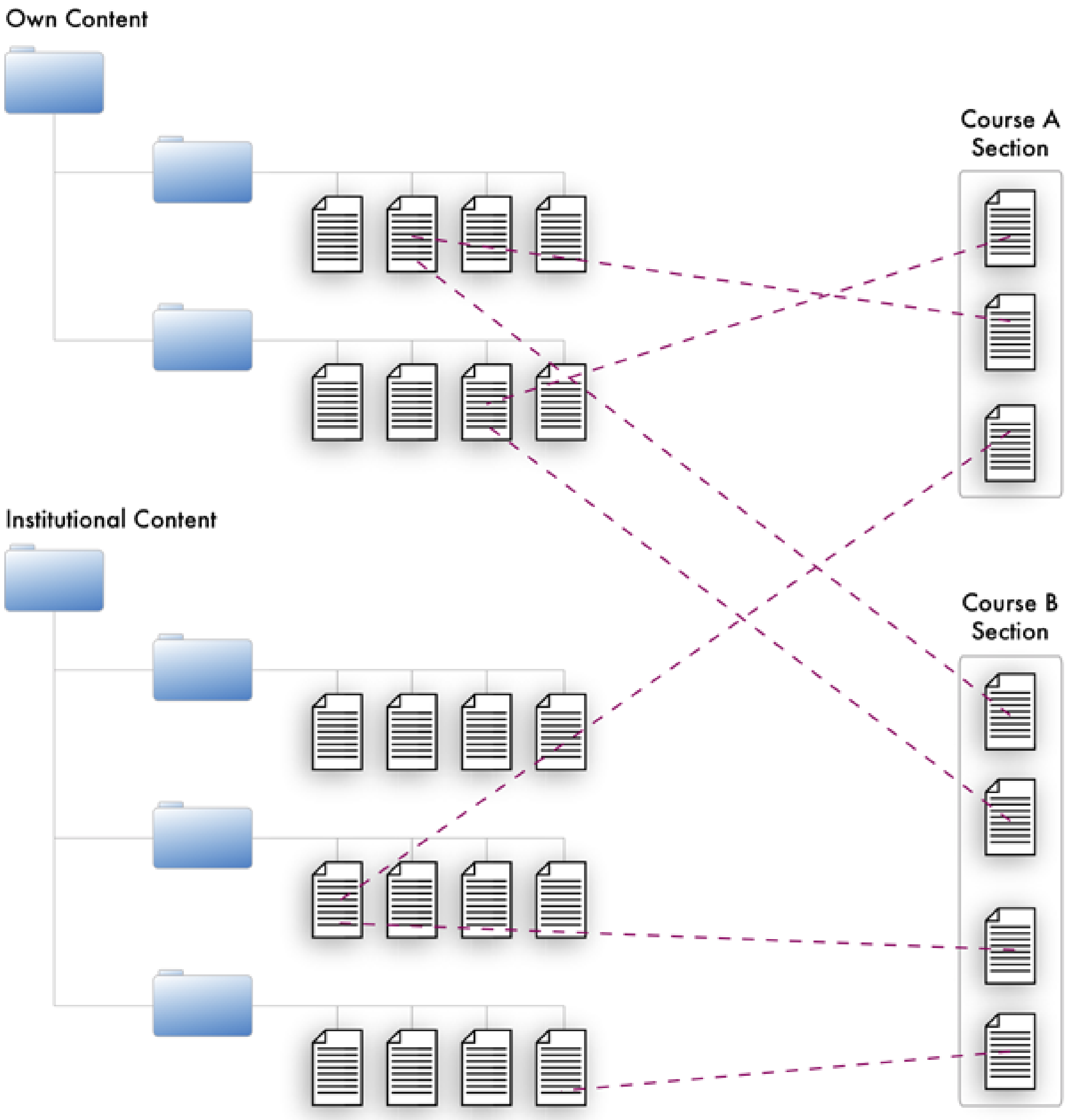



***Figure 4:*** *File-based content organization in Blackboard Content System*

**Blackboard Content System (BCS)**. Blackboard's commercial content management solution is a file-based centralized repository that avoids duplication of content used in different courses (see Fig. 4). As no substantial information can be found in public literature, the following statements are based on personal communication with a Blackboard representative (SCHMIDT, 2010). Blackboard currently does not have a standards solution that allows multiple LMSes to share a single content system. The BCS can support both Web Services and open APIs (e.g., WebDAV) that would allow connections to different LMSes. The exact technology used in the customization would depend on the requirements of the customization project. Connectors are offered that can connect to both Blackboard and Moodle but this option has not been realized yet. All content that is stored in is displayed (rendered) using native desktop applications. For example, a .doc or .docx file would open in Microsoft Word, and .pdf file would open in Adobe Acrobat Reader. Users can either be managed in the BCS or the LMS, depending on the system operator’s choice.

## 7 Summary and Outlook

In this second part of our article we asserted there is demand for federated and distributed learning object repositories. We discussed their specific challenges as opposed to centralized architectures and introduced edu-sharing, a distributed repository that exhibits some outstanding features. edu-sharing includes a powerful enterprise content management system at its core, a distributed architecture that leaves management autonomy to the operating institution, integration of heterogeneous e-learning tools and environments such as authoring tools and learning management systems, effective community support, collaboration tools, and cross-institutional networking facilities, all embedded in an extensible portal.

With universal access to the Internet the concept of **cloud computing** is becoming increasingly important (cf. ARMBRUST ET AL. 2009): According to this concept, IT resources (such as data storage, data processing service, data backup, application software, communication services) are no longer operated by users on their own computers or in an organization's data centers, but are provided flexibly as a service by networked data centers. Accordingly, traditional PCs would disappear from classrooms and workplaces. Information and applications would be accessed from mobile devices and new human-machine interfaces. In the “classroom of the future”, the individual learners interaction with the personal computer will be replaced by interaction with shared digital equipment (e.g., an interactive whiteboard) or via mobile devices (e.g., smartphones and netbooks) while information and communication is processed by efficient services based on Internet technologies. In this way, advanced management systems for digital learning content can form the technical basis in the background for dynamization of learning with mobile devices.

In this context, the trend toward networked repositories for learning content and learning arrangements suggests that the existing monolithically structured virtual learning environments should be replaced by **service-oriented architectures**. Functions of learning management systems that are not part of the core management of teaching and learning processes (such as services for communication and cooperation in web forums and Wikis) will be replaced by specialized systems, which can be orchestrated in teaching and learning processes as required (WILSON 2007). This trend is apparent in exercise and test systems that are replacing the test functionality of virtual learning environments. In addition, repositories for learning content are replacing the comparatively limited functionalities of the learning (content) management systems. Current learning (content) management systems will become specialized systems for the pedagogical arrangement of teaching and learning processes, which ideally will make tools, services, and learning materials available in different systems useable in various educational settings.

At the same time, access to the specific functions used in e-learning will no longer be limited to the course room of a virtual learning environment. Educational institutions, especially universities, are increasingly using (web-based) **portal solutions**. These aggregate different systems, services, and information products used in studying (teaching and research) in a role-specific or individual manner. Similar to this is the integration of products and services for knowledge transfer in companies and knowledge management in personalized portals for the employees. Learning object repositories provide some of these specialized functions in personalized study portals and personal learning environments (PLEs). They offer core functionalities of management, accessibility, controlled access, and use. Via context services, they also allow relationships between people, groups of people, and content to be surveyed and evaluated as dynamic metadata.

## Acknowledgements

This article is an extended version of a German-language version published in **Handbuch E-Learning, 33rd update, July 2010**:

http://shop.wolterskluwer.de/wkd/shop/shop,1/handbuch-e-learning,
978-3-87156-298-3,deutscher-wirtschaftsdienst,,455/ . It was translated by permission of Deutscher Wirtschaftsdienst, an imprint of Wolters Kluwer Deutschland.

The Deutsche Forschungsgemeinschaft (DFG, the German Research Foundation) sponsored the R&D work of the project CampusContent, which laid the methodological and technical foundations of edu-sharing, under code number 44200719. The authors would also like to thank all collaborators and students who were involved in the design, implementation, and evaluation of edu-sharing. Finally we thank the anonymous reviewers and Mike Smit for valuable suggestions that helped us to improve this article substantially.

## References

ALFRESCO: Open source software for enterprise content management, 2005, http://www.alfresco.com/ (last check: 2010-11-24)

ARMBRUST, MICHAEL; FOX, ARMANDO; GRIFFITH, REAN; JOSEPH, ANTHONY D.; KATZ, RANDY; KONWINSKI, ANDY; LEE, GUNHO; PATTERSON, DAVID; RABKIN, ARIEL; STOICA, ION; ZAHARIA, MATEI: Above the Clouds: A Berkeley View of Cloud computing. Technical Report No. UCB/EECS-2009-28, University of California at Berkley, USA, Feb. 10, 2009.

BARANIUK, RICHARD G.: Challenges and Opportunities for the Open Education Movement: A Connexions Case Study. In: Iiyoshi, Toru; Kumar, M. S. Vijay; Brown, John Seely (editor), Opening up Education: The Collective Advancement of Education through Open Technology, Open Content, and Open Knowledge. Cambridge, Massachusetts; London, England: The MIT Press, pp. 229–246, 2008.

CAMPUSCONTENT, 2010, Project Website: http://www.campuscontent.de/ (last check: 2010-11-24)

CHESBROUGH, H.: Open innovation. The new imperative for creating and profiting from technology. Harvard Business School Press, 2003.

CREATIVE COMMONS: Non-profit corporation providing free licenses and other legal tools to mark creative work, 2010, http://creativecommons.org/ (last check: 2010-11-24)

DFG: Deutsche Forschungsgemeinschaft. 2010. http://www.dfg.de (last check: 2010-11-24)

DUVAL, E., FORTE, E., CARDINAELS, K. ET AL.: The Ariadne knowledge pool system, pp. 72-78, Communications of the ACM, Volume 44, Issue 5, 2001

EDMOND: Elektronische Distribution von Medien ON Demand der Medienzentren in NRW. 2010, http://edmond20.lvr.de/ (last check: 2010-11-24)

EDULAP: Project “Educational Landscape Psychology“, Universität Zürich, 2008, http://www.edulap.uzh.ch/ (last check: 2010-11-24)

EDU-SHARING: 2010, Homepage of edu-sharing.net: http://www.edu-sharing.net (last check: 2010-11-24)

GLOBE: Global Learning Objects Brokered Exchange, 2008, http://www.globe-info.org (last check: 2010-11-24)

GOOGLE WEB TOOLKIT: A development toolkit for building and optimizing browser-based applications, 2006, http://code.google.com/webtoolkit/ (last check: 2010-11-24)

Java Community Process: JSR 170 - Content Repository for Java technology API. 2010, http://www.jcp.org/en/jsr/detail?id=170 (last check: 2010-11-24)

KLEBL, MICHAEL; KRÄMER, BERND J.; ZOBEL, ANNETT: From Content to Practice: Sharing Educational Practice. In: edu-sharing, British Journal of Educational Technology, Vol 41, 2010, Number 6, pp. 936-951

KORTEMEYER, GERT: LON-CAPA – An Open-Source Learning Content Management and Assessment System. In: Siemens, George; Fulford, Catherine (editor): Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009. Chesapeake, VA: AACE, 2009, pp. 1515-1520.

KRÄMER, BERND J. & KLEBL, MICHAEL: Supporting Student Mobility and Collaboration through the E-Learning Portal edu-sharing, Fourth Colloquium on International Engineering Education, Research Report 4/2009, Department of Mathematics and Computer Science, FernUniversität, 2009, http://deposit.fernuni-hagen.de/2684/1/Forschungsbericht_4_2009.pdf (last check: 2010-11-24)

KRÄMER, BERND J.; KLEBL, MICHAEL & ZOBEL, ANNETT: Sharing Educational Knowledge and Best Practices in edu-sharing, Second International Conference on Mobile, Hybrid, and On-line Learning 2010. IEEE Computer Society, 2010, pp. 53-59

KRCMAR, HELMUT: Informationsmanagement. Springer, Berlin, 4th edition, 2005.

LEARN:LINE NRW: Die Bildungssuchmaschine des Landes NRW. 2010 http://www.lvr.de/app/LearnLineSuche/ (last check: 2010-11-24)

LIBRARY OF CONGRESS: Search/Retrieve via URL (SRU), 2008. http://www.loc.gov/standards/sru/ (last check: 2010-11-24)

LIFERAY: Open source portal and collaboration software, 2004. http://www.liferay.com/ (last check: 2010-11-24)

MERLOT: Merlot Multimedia Educational Resource for Learning and Online Teaching, 1997, http://www.merlot.org/merlot/index.htm (last check: 2010-11-24)

OPEN ARCHIVES INITIATIVE: Protocol for Metadata Harvesting, 2002. http://www.openarchives.org/OAI/openarchivesprotocol.html (last check: 2010-11-24)

SCHMIDT, ADREAS: Blackboard International BV, Amsterdam, personal communication. 2010

SHIBBOLETH®: Shibboleth Federating and Single Sign-on Solution, 2010. http://shibboleth.internet2.edu/ (last check: 2010-11-24)

SIMON, BERND; MASSART, DAVID; VAN ASSCHE, FRANS; TERNIER, STEFAAN; DUVAL, ERIK; BRANTNER, STEFAN; OLMEDILLA, DANIEL; ZOLTÁN, MIKLÓS: A Simple Query Interface for Interoperable Learning Repositories. Workshop on Interoperability of Web-Based Educational Systems in conjunction with 14th International World Wide Web Conference (WWW'05). ACM, 2005.

SODIS: Datenbank der deutschen Länder und Österreichs für Medien in der Bildung, 2009. http://www.sodis.de/ (last check: 2010-11-24)

water & stone. Open Source Content Management System Market Share Report, 2010, http://www.waterandstone.com/sites/default/files/2010%20OSCMS%20Report.pdf (last check: 2010-11-24)

WENGER, ETIENNE: Communities of practice: Learning as a social system. The Systems Thinker, 9. 1998, http://www.ewenger.com (last check: 2010-11-24)

WILSON, SCOTT; LIBER, OLEG; JOHNSON, MARK; BEAUVOIR, PHIL; SHARPLES, PAUL & MILLIGAN, COLIN: Personal Learning Environments: Challenging the Dominant Design of Educational Systems, In: Journal of e-Learning and Knowledge Society, 2, 3, 2007. http://je-lks.maieutiche.economia.unitn.it/index.php/Je-LKS/article/view/146/144 (last check: 2010-11-24)

W3C: XML Query (XQuery), 2006, http://www.w3.org/XML/Query/ (last check: 2010-11-24)