Legal Engineering - Introducing legal thoughts to the design of an online learning platform

Michael Sonntag

When designing systems, legal thoughts often come only after completion: The system is already in use and someone raises legal concerns, which are not or not adequately covered. In these late stages of a project changes are very expensive. While some legal requirements (e. g. informing users) can be rather easily added later, most others (e. g. privacy: access by staff or other users of the system) are very hard to add/correct if not initially included. This is even more important, as many issues are rather simple to implement at the beginning, but difficult and/or expensive to add or change later on (e. g. obtaining and storing consent pertaining to certain material). A solution to this problem is adding legal considerations to the conception/design phase of a project. The results from this approach in a project for creating an online learning platform as well as general ideas how to introduce this into software development methodologies are presented.

1. Motivation

An Online Learning Platform (OLP) is a web-based environment for cooperative learning [7] and teaching. It consists of presentation of learning materials, communication facilities and interactive elements. Its main functions are delivery of content and administration of courses. Examples of such OLP's are WeLearn [13], Blackboard [1], or WebCT [12]. WeLearn was created at this institute and in widespread practical use. It is a framework consisting of four major components. The WeLearn learning environment (shortly called WeLearn-system) offering the basic online functionality, the settings for schools, universities and adult continuing education providing ready-made templates for different types of courses, course material (specified and edited especially for DE) and the WeLearn Offline Converter that converts courses in CPS format to an offline (D)HTML or Applet version for providing material offline, e. g. on CD's, in a similar manner as online within the platform. Currently an enhanced version of WeLearn (Codename Emerald) is under development, which will include features like a workflow component, agent integration and personalization for working towards a fifth generation model of distance

education (as described in [10]) by supporting computer mediated communication. The current version, however, suffers from several legal problems like missing consent for using personal data, showing too much information in some places or the lack of functions to remove certain content. These are partly solved outside the system (e. g. voluntary use only) and partly just ignored (as long as nobody complains, there is no problem). This is not a desirable state, but cannot really be changed as this would require a lot of effort. Therefore legal thoughts should have some place in the initial design of the new version of the platform.

In the second chapter, some related work will be presented on how and where legal issues are included in designing software. Chapter three discusses legal issues for OLP in general, while chapter four elaborates in which design phase these issues should be considered. In chapter 5 the WeLearn system is presented briefly, followed by a description of the results of applying legal thoughts to its design in chapter 6. Some conclusions end this paper.

2. Related work

Related work is very scarce to the authors knowledge. There are currently three general positions:

- No considerations: This is not really a solution as it does not deal with the problem. Legal issues must be handled outside the system, using conventional paperwork, oral instructions, etc. or relys on nobody attempting anything illegal. As far as this is insufficient, risks are handled through insurance or not at all. In some areas this is however a very viable solution. E. g. in America privacy laws are much different and very liberal, so precautions required for complying to European law are unnecessary there. Other examples are in the area of spam (opt-out instead of opt-in) or intellectual property rights (e. g. works created by government employees are not copyrightable; Title 17 chapter 1 Sec. 105 US Code [11]). This is a problem when exporting to or using software in other countries.
- Legal or government systems (e. g. record handling or document retention systems): Compliance to the legal requirements is one of the basic functionalities and already included in requirements definition. Often a current (and presumed to be legally correct) handling is just transferred to an electronic way, replacing paper with files. These systems are distinct from those targeted here, as they have no value at all if they do not comply with the rules. We focus on those systems, where

legal regulations are "add-ons" but no core functionality. They are not needed for the application to fulfill its actual work, but should nevertheless be considered.

- Legal thoughts as byproducts: Legal rules are ideally self-evident in the form, that they are clear to anybody without them reading the rules. So sometimes they are followed (or systems are enabled to be compliant) through implementing a good design. Examples are enabling users to employ pseudonyms or define who can see their actions. This is implemented as a feature and as a solution if someone has concerns, but not because of legal analysis. Other examples are avoiding any logging because of leaving this kind of functionality out because of a lack of time. Another example is sending E-Mails only in response to registrations and perhaps sending a confirmation mail before actually placing an address on a list. The last is however a tricky example: Opting-in might require active actions by the user (see below), so a pre-selected checkbox might seem sufficient from a technical point of view but not if seen legally.

3. Legal requirements of OLP

Legal requirements result from different sources like public and private law, European and national law, and can also be categorized according to their topical content. All areas mentioned below (with the exception of the teaching laws) are of importance for any collaborative system. Additional ones might be required for the topical area (e. g. in E-Commerce or for medical doctors). The important ones when talking about distance education are the following:

- Privacy: Not everyone is allowed to see every bit of information about others, rather the contrary. This is however a special legal requirement as with consent of the affected person almost everything is allowed. Obtaining this consent is not that easy however, and requires careful design. The user must be exactly informed which data is collected for which use and who will (or will not) receive it. This is necessary, as no general consent is possible, but only for special cases[1]. These need not detail every single use, but must at least list the broader categories. Also the implementation is important: Either e. g. a checkbox for giving permission must be unchecked (so the user has to take some active action to enable using his/her data) or it must be especially visible (larger font, strong color, etc.). This is e. g. even stronger in Germany, where conscious action is required (§ 4 para 3 TDDSG). Initially even a digital signature was required, but this was

[1] E. g. giving consent to the use of the personal data "by all companies of this group" is insufficient information [5], the individual companies to receive the personal data should have been listed (accompanied with what they will do with it).

later abolished. Practical issues in an OLP are awareness (seeing who is online and working where, etc.), shared workspaces (access to other's documents), forums (showing additional information like IP-address; or who may access it, either initially or later on), and access to all data through administrative staff, which must then be under an obligation of secrecy.

- E-Commerce laws: Special laws for E-Commerce (E-Commerce directive and its national implementations [3]) require informing users. They do not apply to OLP, as they are mandatory only for commercial services (provided for remuneration). However, free education is a university-exception; commercial providers do exist. One of the requirements is information about the service provider (e. g. name, geographic address, trade register number, etc.). This is rather easy to comply with, but must not be forgotten. Applicable also only to commercial services but important also for free portals is the liability of its provider (Art. 14). An OLP is in the category with strictest regulation: hosting provider (not for its own content, it is always liable for this, but for content from users, e. g. postings in forums, comments on material). Important is, that liability is independent of the publicity of the information: Private comments to material viewable by nobody else are included. One of the requirements is that the provider must be able to remove the material or at least disable access to it, with the latter being the preferred. Therefore an OLP must contain methods and procedures for disabling access, while leaving the content in place (no actual deletion). This must be done in response to complaints expeditiously. This especially applies also to the owner of the material, so the system for controlling access must also allow disabling this type. Also, general procedures must be set up: Who will have to do exactly what?
- Laws against SPAM: Part of many regulations is unsolicited commercial communication. Important to note is, that communication is very easily commercial. If it serves to further a commercial service or product in any way (e. g. sending an SMS to notify of winning a prize to be collected by a premium rate telephone call[2]), it must conform to these rules. Automatic notifications about events in the system (e. g. new material available, postings in forums) are an important part of an OLP. Again, if this is a commercial platform, these can be seen as spam if sent without request (they try to further the use of the platform which must be paid for and are therefore commercial). As already a single mail, asking for permission to send is prohibited, on registration (or by popups on webpages) the user must be asked for receiving these notifications. Again a legal but

[2] UVS Steiermark, notification 29.3.2002, GZ 30.2-153/2001; See also VfGH decision 10.10.2002, G267/01, G268/01 and several others. For Germany e. g. LG München I 15.4.2003, 33 O 5791/03

not technical requirement (storing the fact would suffice) is, that time and circumstances of this consent must be recorded (logs). Obviously, possibilities for unregistering must also be offered.

- Consumer protection: Consumer protection laws vary largely in different countries. They are again only pertaining to commercial OLPs. However, these requirements are very general and computer or communication specific. Examples are content and presentation of general terms of contract, requiring archiving of previous versions, or the obligation and period to inform about rights for rescinding a contract.
- Computer crime: No direct requirements stem from this source, but here again methods for blocking access should be included to avoid aiding in an offence after notification. Important to note is, that no general obligation for monitoring users is required or allowed, so no special precautions (automatic checking, marking the time of the last visit, etc.) are necessary. However, extensive logging of user's actions should be available (but not activated by default because of privacy issues; see also above) to be able to fulfill demands of courts for securing evidence. This must be severely restricted on the other hand, as listening to private communication is expressly forbidden (in the form of mail, applying also to E-Mail, but not e. g. newsgroups), so access to logs should only be available to a very small group. Seen from the other side, certain precautions should be taken to be protected by these laws, e. g. working security measures within the system (not external to it[3]) for accessing the platform or the material (especially downloadable one).
- Intellectual Property Rights: Important parts for OLP, both for teaching content (course material) as well as content from learners are intellectual property rights (IPR). Even if work is handed in by students for examination, it is still copyrighted and cannot just be reused in a different context (e. g. distributed on Offline CD's, used as material for the next year's courses, ...) or published on the Internet. Because of this, acquiring the rights from participants (which they might or might not transfer) is necessary. This information must be stored (e. g. in metadata associated to documents or folders). For example when exporting the data or copying it to a new location (e. g. when creating a new course), this information must be shown to the user, as automatic checking is not really possible. Also, reusing published content (from other sources, e. g. the internet) is allowed for non-commercial teaching (special exception), but only applies to certain forms or numbers (only to participants of the course). This requires granting access to it only to students

[3] A system with its content is only protected against intrusions subverting security measures within this system. Therefore cookies are insufficient for recognizing users: Their target is recognition but not security (as with passwords). They should only be used with a short expiration date and another means of identification (e. g. passwords).

taking this course but not to the public (which might be allowed to view other parts of the course). For commercial OLP, rights for any content must be acquired if it is to be used. Then again marking with metadata is needed to be able to restrict use to those cases allowed by the license. Another IPR-issue is the protection of metadata. Metadata used for electronic rights management may not be removed or changed, therefore it must be stored unaltered (no stripping away) and additions are only allowed insofar they do not change the original meaning.

- Teaching laws (UOG): In Germany a special Distance Education law (Fernunterrichtsgesetz [5]) exists. Its main intention is consumer protection and no direct consequences for OLP result from it[4]. Interesting in connection with privacy might be that the organizer of a course is required to supervise the learning progress, resulting in a legal basis for continued observation (which can not, however, replace consent as other methods of supervision are also possible). No similar law exists in Austria.

4. Legal issues for the different design phases

In this part the issues previously mentioned are allocated to the different stages in design. For this, the waterfall and the V-model are used as basis. Iterative models (like the spiral model) are much more difficult in this respect, as no clear allocation is possible. With respect to legal requirements they are kind of "dangerous", as these are sometimes important on many levels (from general design down to implementation, making them difficult to add later), usually possess rather low priority (if included at all) and rarely show in UI's or provide functionality to end-users and are therefore easily overlooked e. g. when using rapid prototyping. Coding and testing phases are important, but excluded here, as legal requirements are identical to any other requirements defined by users or technical reasons in these phases. The only difference might be that sometimes the absence (e. g. personal information) must be tested, which is rather hard, compared to testing something being in place.

4.1. System requirements formulation

In this phase the functions the system should fulfill at the end are specified, respectively the actual wishes of the customer should be clarified.

[4] It requires (§ 12) e. g. prior approval by administration for commercial distance education courses, but only if the course (content, organization) is not suitable for reaching the claimed educational level.

One aspect to take care of here is privacy: Which data is actually needed for fulfilling the basic functionality and therefore indispensable, which is required for additional "goodies" and which is not really needed (at least now)? How and from which source (the person itself, derived data, from external sources, …) is it gathered? What are the legal justifications for each of them, so where is consent needed (and must it be explicit or is implicit consent sufficient)? What will happen with the data later on (will it be reused, published, made available to certain groups, etc.) and to whom will it be disclosed? All this determines both the privacy policy, as well as the implementation. A connected issue is awareness, as it determines partly the use and disclosure of data.

Methods for blocking access and deleting content must be included to fulfill requirements of E-Commerce and criminal laws. Procedures for who is allowed (respectively required) to look at private content must also be established and supported by software.

With respect to IPR, the expected categories (producers) of content data must be analyzed similar to privacy issues for person-specific data. What metadata will be required for marking the creator and the allowed usage of the material? Whether and how will this be viewable/exportable and be enforced within the system?

Another issue is SPAM and notifications. What notifications (content) should be offered and how often will they be sent? Which method of delivery will they use?

4.2. System design

The global functions of the software and the partitioning in modules (or objects, agents, etc. according to the used design paradigm) are described in this phase.

With respect to legal issues, the areas identified in the previous step must be added to the individual modules. Probably very few special "legal module" exists, but rather legal extensions to others. To a large extent determined by law and possessing dedicated functionality large enough to be designed as an independent part are the areas of privacy (obtaining and storing consent for different topics, import and export functionality allowing to select parts and mask fields to transfer only the necessary minimum) and intellectual property rights (entry/import of metadata, how and where to store it, the format to use,…).

Also the rights system (who may access which elements) must be detailed here, taking care to set the default permissions correctly. For example administrative staff should not possess global access by default, but rather be required to explicitly obtain it individually (probably not for each item separately to ensure smooth working).

4.3. Logical design

The interactions between the independent units of the software are specified here.

Examples are how modifications to IPR protected data are made or prevented as well as how access is verified. Other issues are exactly how and when information is shown to users and how they can react: Privacy requires informational messages to be shown several times (e. g. before the first use, the first time after changing the privacy policy) or in several places (e. g. on the first use of special functionality). The triggering mechanisms for this as well as the resulting actions when the user declined must be specified here. Also important is e. g. the internal transfer of data. Personal data must be protected so communication between different parts of the system must be encrypted of wiretapping is possible (e. g. when storing the data in an external database).

4.4. Unit design

In the detailed design of individual parts legal requirements do not come into play directly any more but as influences, e. g. with exactly which text and in which state (e. g. checkboxes initially checked or not?) forms are presented to users.

Examples for this phase are the actual text and presentation of information which must be shown to the user and in which way he/she must acknowledge this. This is important for many parts and issues (privacy, notifications, obtaining consent, correcting mistakes, …) and should therefore be consistent throughout the system.

Another example is metadata for the protection of IPR. This phase must specify which parts of the metadata and what of the related content is shown for which person and what is suppressed. This is important e. g. when showing content within a webbrowser, as there metadata can be embedded, but is usually not shown. So precautions both for saving (including it in the document) as well as printing (including it in a readable version) are necessary.

5. The WeLearn platform

WeLearn (Web Environment for Learning, [13]) is an online learning platform developed at this institute. It is, as already mentioned, a framework consisting of 4 major components: The WeLearn learning environment offering the basic functionality, the different settings for schools, universities and adult con-

tinuing education (providing ready-made templates for different types of courses), course material created especially for DE (as examples and for reuse) and the WeLearn Offline Converter for converting courses in CPS [6] format to an offline (D)HTML and applet version (to allow distribution on CDs).

The WeLearn-system is a free and open Distance Teaching/Coaching/Learning environment [2] implemented in Java for platform independence. It is universally useable, which means that using it is not limited to specific target groups, didactical models or training areas. Therefore one major design goal was keeping the system easily and intuitively useable, e. g. through using a browser for presentation and adhering to their common usage (see *Figure 1*). The same is true for the teaching and administration side, which is also intended to be intuitively useable and as simple as possible. The whole development was driven by this idea (providing complexity but hiding it to the utmost degree possible), as well as platform independence and accessibility. WeLearn has reached these goals as feedback from practical use at several institutions (University Linz, University Zurich, BRG Wagrain, HBLA Steyr [8]) shows. Users can access their courses everywhere simply by using a web browser (especially useful for coaches). This and viewers appropriate for the content used (if not HTML; e.g. Acrobat reader) are the only requirements.

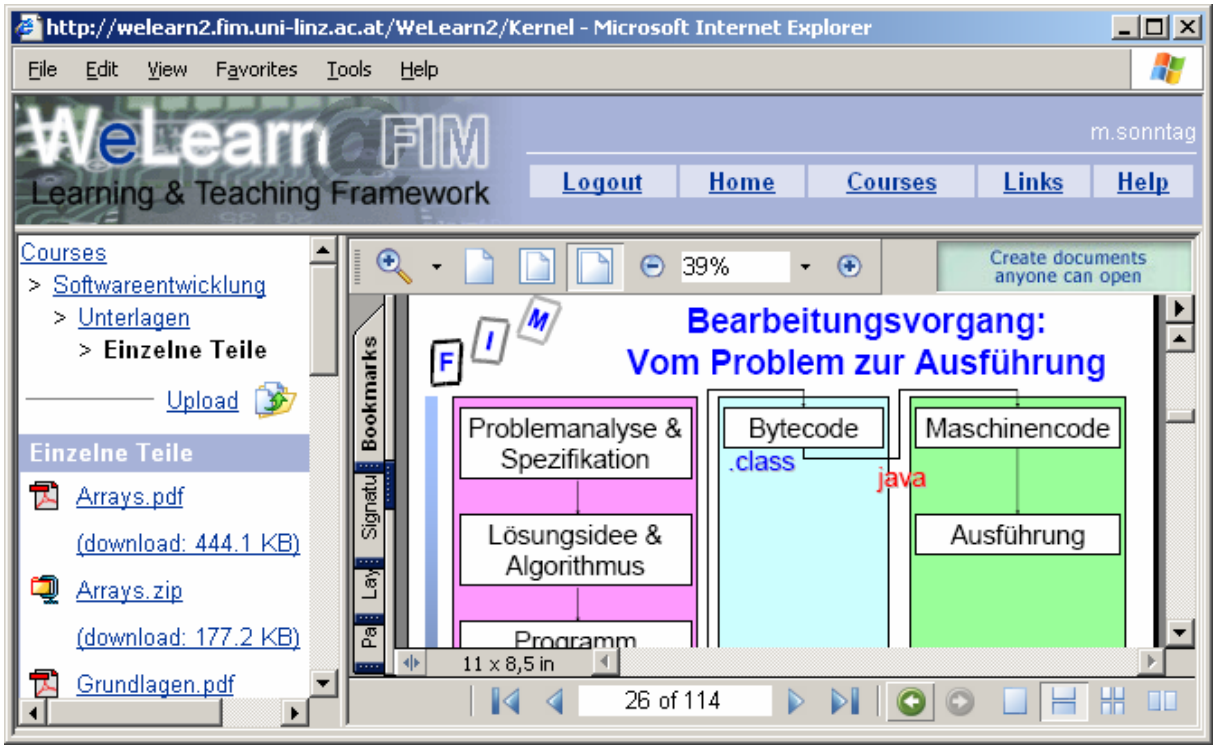

Figure 1: WeLearn with an open course

However, judging from our experiences and the feedback we received from our partners and users (course providers and learners) we are convinced that further improvement of the system is possible. As

an answer to this we are now working on a two-year project for realizing the vision of a personalizable and adaptive platform supported by agents (using the locally developed agent platform, [9]), which will be a redesigned and enhanced version of the existing framework. This project will for the first time include legal design right from the beginning.

6. Practical application: Resulting technical requirements

The project is currently partly in the design phase (advanced functionality like the agent interface) while the basic framework is currently implemented. Therefore no practical experiences but only the design issues identified can be described now. In detail, the following features have been introduced to the system because of legal engineering:

At the first login (and configurable later; can be associated with any folder) a configurable message is shown and user input is requested. A checkbox must be activated to show agreement with these terms before the system (respectively folder) can be entered. This action is stored with the associated data (login name, IP address, time) in the user model in an additive way (later actions, e. g. if additional consent is retrieved for new/changed parts, add to this list but do not replace it). As registration of participants is done by administrative staff (or automatically, e. g. retrieving the data directly from the university database), this must be done then instead of the (missing) registration. This is not possible in the current version of WeLearn at all: Only an information message could be placed, but enforcing showing or having to acknowledge it is impossible. This is intended for acquiring consent for using personal data (privacy) as well as acquiring IPR for submitted content. A user interface for changing these data (e. g. revoking the consent, which must be possible) is not planned: In such a case the user should be removed from the system. If this is desired later, it can always be easily added by introducing a new object and checks in the special objects dependent on individual consents.

Notifications about new or changed content (e. g. uploads of projects by students) must be explicitly configured for content, frequency and method of delivery. By default, no messages are sent. Also, the preferences can be changed at any time, although only within the system (no unsubscribe-link in E-Mails).

An extensive event system will be included which also supports logging of all events, both for providing functionality (personalization) as well. Because of the huge load this would pose, filters will be implemented to select only interesting events (according to the publish-subscribe principle). Events will also

be included in the security system to allow viewing events only by those authorized to do so (which by default depends on the access to object generating the event).

Every object within the system can be associated with arbitrary metadata (in XML format). This allows preserving additional information even if it is not understood and used by the system, which is a requirement e. g. for rights-management data. Some objects will also interpret parts of the metadata, e. g. course packages (the CPS standard used employs the LOM metadata standard, which is very comprehensive).

Sending private messages will initially not be included as it poses legal problems, e. g. not even administrators would be allowed to take a look at them. This poses problems especially in the initial testing phase and for finding/correcting problems. The alternative, declaring all private messages as public and retrieving consent for accessing them is not a desirable and viable alternative.

7. Conclusions

Legal aspects can be important for both large (obtaining consent) and small issues (what parts of general terms of contract must be in a different color/boldface). The former ones must be included already in the design phase, as they also determine what functionality can be provided or what additions are necessary when offering them. It is therefore important to recognize the additional restrictions right from the start, as adding them later might be difficult. Legal requirements are usually not that complicated for developing the software itself. However, this is a kind of intersectional ascpect, reaching from the design down to the actual implementation.

The most important legal aspects for online learning portals are privacy issues and how to handle content from different sources: external teaching material and those from the teachers, contributions by students and examinations. All must be tagged according to their creator and the type of use allowed.

Some questions and issues were presented to be answered or taken care of during the different stages of designing software to avoid most legal problems for the resulting system. The presented ideas are however, tailored mostly for online learning portals. For other projects, many issues (like privacy) will be similar, but others (e. g. teaching laws, spam or IPR) will be of less or no concern, while yet other ones must be considered. This is complicated by having to comply with potentially many areas and sources (from each country the software is intended for) of law, yet it is unavoidable as law is more and more enforced also in the "virtual" world because of its increasing importance and awareness of problems that can arise there.

8. Acknowledgement

This paper is a result of the project “Integrating Agents into Teleteaching-Webportals“ sponsored by the FWF of Austria (Fund for the support of scientific research; Project number P15947-N04).

9. References

[1] Blackboard: http://www.blackboard.com/

[2] Divotkey, R., Mühlbacher, J.R., Reisinger, S., Remplbauer, D.: The WeLearn Distance Teaching Framework. Granada: Proceedings of the 2002 Eden (European Distance Education Network) Annual Conference 2002

[3] E-Commerce directive: Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce) OJ 17.7.200 L 178/1-16

[4] "Friends of Merkur" decision by the Supreme Court: OGH 27.1.1999, 7 Ob 170/98w

[5] German Distance Education Law: Gesetz zum Schutz der Teilnehmer am Fernunterricht (Fernunterrichtsschutzgesetz - FernUSG) from 4.12.2000 (BGBl. I S. 1670), last changed 23.7.2002 (BGBl. I S. 2850): http://www.bmbf.de/pub/fernusg_neu_2002.pdf

[6] IMS Content Packaging Specification, Version 1.1.2 Final Specification: http://www.imsproject.org

[7] Jechle, T.: Tele-Lernen in der wissenschaftlichen Weiterbildung. In: Dittler, U. (Ed.): E-Learning. Erfolgsfaktoren und Einsatzkonzepte mit interaktiven Medien. München: Oldenburg 2002

[8] Mühlbacher, J. R., Mühlbacher, S. C., Reisinger, S.: Learning Arrangements and Settings for Distance Teaching / Coaching / Learning: Best Practice Report. In:Hofer, C.,Chroust, G. (Eds.): IDIMT - 2002; 10th Interdisciplinary Information Management Talks. Linz: Universitätsverlag Rudolf Trauner 2002

[9] Sonntag, M.: POND - Ein Agentensystem mit Fokus auf Sicherheit und Bezahlung für Leistungen unter Berücksichtigung rechtlicher Aspekte (POND - An agentsystem with focus on security and payment for services with consideration of legal aspects). Dissertation. Johannes Kepler University Linz: 2002

[10] Taylor, J. C.: Automating e-Learning: The Higher Education Revolution. In: Schubert, S., Reusch, B., Jesse, N. (Eds.): Informatik bewegt. Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.v. (GI).Bonn: Gesellschaft für Informatik 2002

[11] US Code: http://www4.law.cornell.edu/uscode/

[12] WebCT: http://www.webct.com/

[13] WeLearn: Web Environment for Learning: http://www.fim.uni-linz.ac.at/research/welearn/