RE-USABLE LEARNING OBJECTS IN PRACTICAL USE

Gerd Kortemeyer (kortemey@msu.edu)
Michigan State University, East Lansing, MI 48825, USA[*]

Keywords: Learning Objects, Content Sharing

ABSTRACT

The idea of re-usable learning content objects has been around for while, but not many studies exist showing how much content sharing is actually happening. The paper explains the design principles of one particular content sharing network, namely the Learning*Online* Network with CAPA (LON-CAPA), and provides statistics regarding the actual flow of online resources between institutions.

INTRODUCTION

As educational institutions establish an online presence, initial successes are often due to individual faculty members ("early adopters" of this new technology), working long hours to develop material more or less single-handedly. Frequently, they are leaving behind scattered projects, which are of intrinsic value, but of little use for the institution and far less for the larger academic community. The same is frequently true for content developed in externally funding curriculum development projects, where stewardship for the materials oftentimes ends with the end of the funding period, and little or no sustainable dissemination strategy is in place. “Late adopters” of technology in education might altogether refuse to venture into creating new online educational resources, since the task of creating comprehensive material appears overwhelming in isolation. To address these problems, an infrastructure to provide a course and learning content management system was created, which has resource sharing at the base of its architecture: the Learning*Online* Network with Computer-Assisted Personalized Approach (LON-CAPA).

The roots of this system go back to 1992, when a group of faculty at Michigan State University started developing a sophisticated online homework and assessment system, with a strong focus on the sciences and mathematics[1]. Soon other universities adopted the system, and it was not long before an informal culture of inter-institutional sharing of such resources developed. The exchange of resources however was reflected nowhere in the architecture of the system, and was achieved by copying files (at the time, on floppy disks), FTPing them from server to server, or emailing them from colleague to colleague. To formalize and thus further the content sharing within the emerging community of practice, the team in 1999 started to add cross-institutional digital library and learning content management capabilities, which were not limited to problem resources, but include the management of any multimedia materials (images, HTML pages, applets, etc)[2].

The top layer of the system consists of a complete course management system, so that the assembled resources can immediately be used in a course context[3], where instructors have the ability to seamlessly assemble their course from resources across institutional boundaries. We believe the course management layer to be essential to provide convenient reuse, including navigation between the elements, provision of secure access control, and avoidance of the cumbersomeness (both technical and legal) of first exporting the resources from some repository or digital library, and then re-importing them into a course management system. The course management system as runtime environment finally guarantees that every resource not only runs, but behaves the same at any participating institution, thus eliminating compatibility concerns.

Today (Fall 2006), LON-CAPA spans 37 universities, 4 community colleges, 44 high schools, 3 middle schools, 6 content development projects, and 6 publishing companies. LON-CAPA is serving over 11,000 students per semester at MSU alone, and well over 30,000 students per semester system-wide.

Its shared resource pool currently holds around 247,000 resources, among those approximately 80,000 original homework and exam problems, 88,000 images, 700 movie and sound files, 1,500 animations and simulations, 58,000 content pages, and 7,400 re-usable content assemblies. Disciplines include astronomy, biology, business, chemistry, civil engineering, computer

[*] currently guest at the CampusContent and CampusSource projects, Fern-Universität in Hagen, Germany.

science, family and child ecology, geology, human food and nutrition, human medicine, mathematics, medical technology, physics, and psychology.

NETWORK INFRASTRUCTURE

The Learning*Online* Network with CAPA is a geographically distributed network of persistently connected servers at schools, colleges, and universities. Each participating institution needs to contribute at least one server to the network (Figure 1). An institution can set up any number of servers within their domain to scale with increasing workload. The network is set up to be redundant and load-balancing.

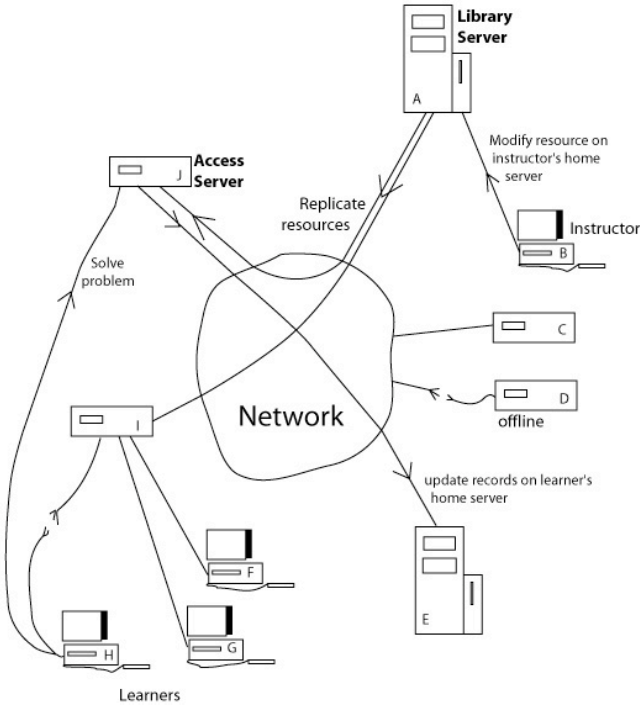


FIGURE 1: THE DISTRIBUTED LON-CAPA NETWORK INFRASTRUCTURE

The network is logically divided into so-called domains, which usually correspond to one institution, such as Michigan State University, North Dakota State University, or Truckee Meadows Community College. Domains can be used to limit the flow of content and the extent of user privileges. Users can log into any server in the network. For example, an MSU user can log into a server at North Dakota State University using his MSU credentials, and find the exact same environment as on one of the on-campus servers.

To faculty users, the distributed content resource pool of LON-CAPA appears as one large virtual file system, and every resource has a system-wide unique and persistent URL path, under which it can be accessed from any server in the network. As users are browsing this filesystem, they are actually transparently accessing content from servers across the network (Figure 2).

As resources are accessed, the network provides resource replication to provide faster access to the resources. When a resource is first accessed on a server in the network, it is copied over in the background, and a so-called subscription is left behind. If the resource is updated on the original server (i.e., a new version is generated), all subscribing servers are notified. If the resource has been used on the subscribing server within a configurable amount of time, it is automatically updated. If on the other hand the resource has not been used for a while, the local copy is deleted, and the server un-subscribes from the resource. The cycle can then begin again should be resource be accessed again on that server at a later point in time.

Instructors of courses using the modified resource are automatically notified of the changes, and can choose to accept the change or stay with an earlier version.

LON-CAPA provides resource versioning and digital rights management. For each resource, it can be determined if it should be public (i.e., available even without logging in), available to instructors system-wide (and if they select it, also to their students), or available only to instructors within the originating institution. Also, the rights for other authors making source-code copies of the resources or modifying the original resource can be configured. Finally, all of these rights can be delegated to access control files external to the resources, so that whole libraries of materials can be selectively opened up or closed to particular institutions or courses.

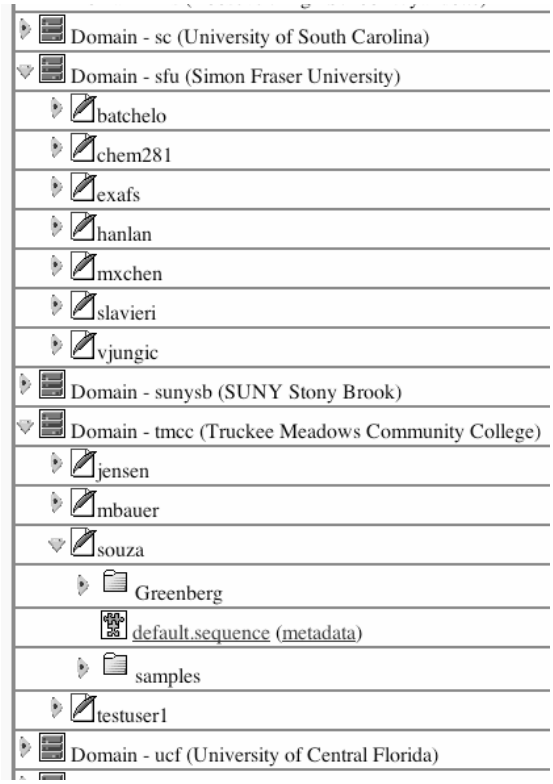
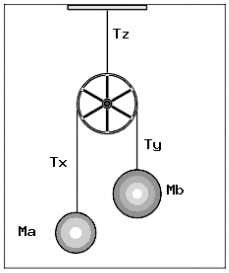


FIGURE 2: THE DISTRIBUTED FILE SYSTEM BROWSER INTERFACE

RESOURCE ASSEMBLY

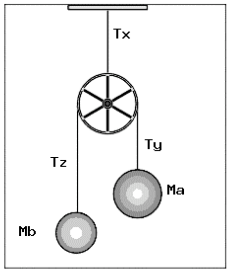
There are different levels of resource granularity in LON-CAPA. At the bottom level are individual resources, also referred to as fragments: simple paragraphs of text, movies, applets, homework problems, HTML pages, etc. Figure 3 shows an example for such a fragment level resource, namely a randomizing homework problem, as seen by two different students (this randomization feature provided one of the initial incentives for joining the network, since it prevents rote copying of answers between students).



A frictionless, massless pulley is attached to the ceiling, in a gravity field g = 9.81 m/s^2. Mass Mb is greater than mass Ma. The tensions Tx, Ty, Tz, and the constant g are magnitudes. (Select a response for each statement.) Motion of Masses on a Pulley.

- Tx+Ty is .... Tz
- Ty is .... Mb*g.
- Ma*g + Mb*g is .... Tz
- The center-of-mass of Ma and Mb does not accelerate.
- The magnitude of the acceleration of Ma is .... that of Mb
- Ty is .... Tx

Submit Answer
Tries 0/2



A frictionless, massless pulley is attached to the ceiling, in a gravity field g = 9.81 m/s^2. Mass Ma is greater than mass Mb. The tensions Tx, Ty, Tz, and the constant g are magnitudes. (Select a response for each statement.) Motion of Masses on a Pulley.

- The center-of-mass of Mb and Ma does not accelerate.
- Tx is .... Mb*g + Ma*g.
- The magnitude of the acceleration of Mb is .... that of Ma
- Tz+Ty is .... Tx
- Mb*g is .... Tz
- Ty is .... Tz

Submit Answer
Tries 0/2

FIGURE 3: TWO VERSIONS OF THE SAME RANDOMIZING PROBLEM

There are several ways of assembling resources into objects at a higher level of granularity: The easiest way is faculty directly assembling fragments into courses. Courses can later be copied ("cloned") for reuse as a whole, and then modified after cloning. Another way of assembling resources is into reusable content assemblies: each time resources are assembled into a higher granularity object, a new reusable learning object is created. For example, an instructor from Community College A can combine a text paragraph from University B with a movie from College C and an online homework problem from Publisher D, to form one page. Another instructor from High School E can take that page from Community College A and combine it with other pages or fragments into a module, unit or chapter. Those in turn can be combined into online course packs (Figure 4).

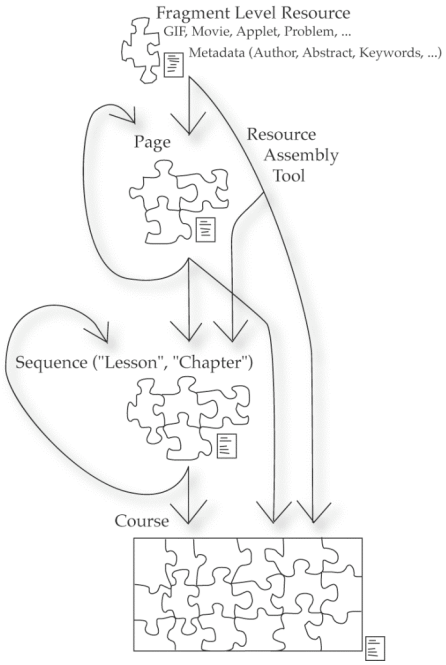


FIGURE 4: RESOURCE ASSEMBLY AT DIFFERENT LEVELS OF CONTENT GRANULARITY

Nomenclature of what constitutes a lesson, a chapter, or a module, is intentionally left ambiguous, since – as it turned out – authors cannot agree on these terms. Also, there are no restrictions on what can be included into what, or how many levels of nesting are allowed or expected.

Instructors can specify the actual path through the learning resources through combinations of learner choices and system-generated adaptations (for example, if the learner does not pass a test, additional resources may be included). Each learner can have an individualized curriculum according to preferences, capabilities and skills. Figure 5 shows a screen shot of the "Resource Assembly Tool" of LON-CAPA, which is used to combine learning resources. In this particular example, each box represents a resource – a single page, or a whole module or chapter. The arrows represent possible paths through the material, and the box labeled "COND" represents a condition under which the path is available.

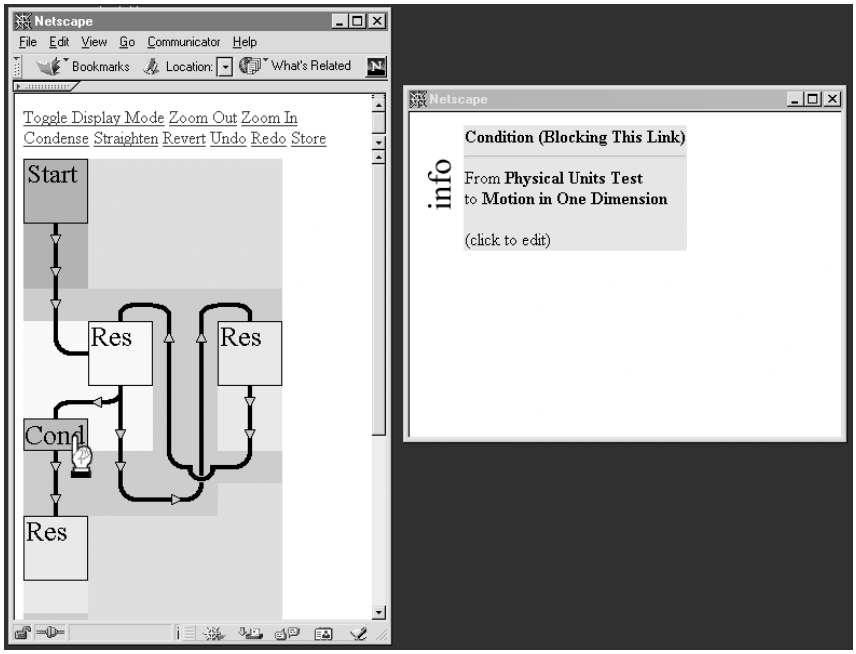


FIGURE 5: SCREENSHOT OF THE RESOURCE ASSEMBLY TOOL

FACULTY AS AUTHORS

Faculty members have written most of the resources in the network, sometimes in connection with grant projects, but mostly originally for use in their own courses. We found a large willingness of faculty to make their materials available to the pool, and the vast majority of the material in the system is published "system-wide." The few exceptions of authors restricting their materials to their own institutions are oftentimes due to hesitation of submitting their material to public scrutiny rather than general unwillingness to share.

At the beginning of the project, we believed that monetary incentives would be needed to accomplish this, but in the meantime, we found that faculty are far less interested in those than in the feeling of accomplishment when they see the usage counters clicking up or are receiving positive feedback from their peers. Proposed payment and bartering schemes were also seen as too complicated, and in the end it was decided that they might inhibit rather than foster the expansion of the network.

It seems that it fits into the academic culture to share resources, just like one would research papers. It should be noted, though, that this might be subject area specific: most participating faculty are from the natural sciences, and most resources are intended for introductory courses. A faculty member may take pride in having written a high quality homework regarding angular momentum conservation, but may hardly base their reputation on it, or feel in competition with peers teaching the same topic. Also, the topics of these courses are far from controversial.

The most important aspect however appears to be good stewardship of the material: the project needs to guarantee that materials, some of them exam or grading-relevant homework problems, do not "leak" out of the pool, i.e., that students only

have access to the material after faculty selected it and opened it up. Particularly sensitive is the XML source code of problems, since it allows for reverse-engineering of the randomization and determination of the correct solution for any variation of the problem.

COMMERCIAL PUBLISHING COMPANIES

The role of commercial publishing companies in the network is somewhat unique. In the US, most courses have mandatory textbooks, which every student in the course needs to purchase at an average cost of more than $100. Most science textbooks include "back-of-the-chapter" problems, which faculty assign. However, particularly large enrollment courses frequently do not have enough assistants to actually grade the student work.

The main business of the participating publishers is the sale of these textbooks, and in order to give faculty an incentive to adopt their book rather than a competing title, the companies offer free online back-of-the-chapter homework problems in connection with adoption of their text. Faculty can thus assign problems from the book and have them graded online. The publishing companies are paying commercial companies to prepare the homework libraries in LON-CAPA format, hosting them, and to selectively open them up to courses using the respective book. In the end, the students are likely paying indirectly through the potentially higher cost of the already expensive book.

Since LON-CAPA is a distributed network, it is sufficient to have these homework libraries hosted on one server in the network. The institutions using the problems can deploy them from their own servers in their own courses, using the mechanisms described above.

As faculty do not get paid for their content (except, potentially, by funding agencies in the framework of curriculum development efforts), and the publishing companies only pay for the preparation and hosting of their materials, there is no actual money flow within the system.

RESOURCE-SHARING IN NUMBERS

With all the tools in place to share content across multiple institutions and deploy them, and the majority of the content published such that it can be reused anywhere across the network, the real question is: does it happen?

LON-CAPA tracks resource usage: every time a resource gets incorporated into a learning object of larger granularity, every time it gets deployed in a course, and every time a learner accesses it, the transaction is recorded and added to the metadata for the resource. The former two events constitute a form of peer-review, with the approval being the adoption of a resource. In addition, data such as degree of discrimination, degree of difficulty, and average number of attempts until mastery of each resource are recorded. Instructors can use this data to make informed decisions when selecting resources for their courses and students, but the same data can be used to track resource flow within the network.

Table 1 shows content provision and usage statistics by resource kind, i.e. the type of multimedia content. The category "other" includes a large variety of media and file types, such as application-specific files, PDF files, raw data, etc.

	Available	Used	Used externally	Instances of being used at institutions
Images	88079	18262	9629	43191
Problems	80009	57979	23816	111764
HTML content pages	57972	5888	2949	15877
Reusable content assemblies	7405	3184	1299	6228
Animations and simulations	1575	507	338	1728
Movie and sound files	742	326	105	587
Other (PDF, MS Office, etc)	11302	2484	674	3693
Total	247084	88630	38810	183068

TABLE 1: TOTAL NUMBER OF RESOURCES AVAILABLE BY RESOURCE TYPE (FIRST COLUMN), NUMBER OF RESOURCES USED (SECOND COLUMN), NUMBER OF RESOURCES USED OUTSIDE THE ORIGINATING INSTITUTION (THIRD COLUMN), AND NUMBER OF INSTANCES RESOURCES WERE USED AT INSTITUTIONS (FOURTH COLUMN).

Of the 247,000 resources available in the pool (listed in the first column), a little more than one third (36 percent) are in active use, where active use means being used recently in at least one course at the participating institutions (second column). Of those, 44 percent are used in at least one course at an institution other than the institution where it originated ("used externally," third column). In other words, 15 percent of the available resources in the pool are shared across institutional boundaries.

These percentages vary strongly by resource type: the percent of active use is the highest for problems, where 72 percent of the available resources are used in courses, and of those, 41 percent at across institutional boundaries – 30 percent of the available ones. Active use is the lowest for HTML content pages. Images are mostly reused in connection with problems or HTML pages – on the average 0.64 images per problem/page.

The last column of Table 1 shows instances of reuse: if the same resource is used at two institutions, it is counted twice. The column is an indicator of how widely the respective media types are distributed across the network. Most resources that are in active use are on the average used at 2.07 institutions, but the spectrum ranges from "other" files, which on the average are used at 1.49 institutions, to animations and simulations, which are used on the average at 3.41 institutions.

Table 2 shows the top dozen resource providers in its columns, and the top dozen resource users its rows. We are not showing the plain text name of each LON-CAPA domain, instead we numbered the universities, community colleges, high schools, middle schools, projects, and publishing companies, and preceded their number with U, CC, HS, MS, PR, and PU, respectively, to indicate the class of the domain. However, it is probably obvious that U01 is Michigan State University.

		U01	U04	PR01	U06	U17	U05	U03	HS20	U12	PR06	U11	U08
	Available	144418	17545	10809	8799	7635	7037	5120	4439	4066	3750	3283	2989
	Used	38245	7596	340	4821	2908	4880	3411	3842	2841	1502	1231	2102
	Used externally	17099	1804	339	974	276	3507	1735	1035	1997	1502	415	62
	Using↓												
U01	38855	*34790*	301	105	17	49	1621	294	74	102	298	137	3
U05	11668	4881	23	14	3	33	*4357*	866	29	500	328	5	3
U04	10343	2393	*6969*		10		207	374	8	128	2	18	
U06	10089	2261	64	13	*4755*		305	1001	8	10	2	72	2
U03	9973	4053	58	27	5	84	1213	*3173*	7	728	14	166	
U08	8578	2014	1078	6	2	2	720	5					*2097*
HS20	6465	2138	1	47			40	350	*3767*	21	70	4	
CC04	6356	1156	25		2	31	1586	789	197	1522		64	7
U17	6270	2689	4	7		*2813*	188	205	94	140	4		2
HS40	5251	3899	22	5		40	65	293	388	70	27	16	1
U14	5135	1682	213	42	12	1	665	42		3	7	114	
U09	4246	3409	7		1			15		1		1	

TABLE 2: TOP DOZEN LEARNING CONTENT PROVIDERS (COLUMNS) AND USERS (ROWS). UNIVERSITIES ARE LABELED "U," COMMUNITY COLLEGES "CC," HIGH SCHOOLS "HS," MIDDLE SCHOOLS "MS," PROJECTS "PR," AND PUBLISHING COMPANIES "PU." THE FIRST ROW SHOWS THE TOTAL NUMBER OF RESOURCES WHICH THE DOMAIN IS MAKING AVAILABLE TO THE POOL, THE SECOND ROW THE NUMBER OF RESOURCES ACTUALLY BEING USED, AND THE THIRD ROW THE NUMBER OF RESOURCES WHICH ARE USED AT A DOMAIN OTHER THAN ITS OWN. THE FIRST COLUMN SHOWS THE TOTAL NUMBER OF RESOURCES A DOMAIN IS USING FOR ITS COURSES, AND THE REMAINING CELLS HOW MANY RESOURCES THE DOMAIN IN THE ROW IS USING THAT ORIGINATED IN THE DOMAIN IN THE COLUMN. ITALICS INDICATE AN INSTITUTION USING ITS OWN RESOURCES.

The majority of resources is both provided and used by Michigan State University (U01). The second largest pool of resources is provided by a much smaller public university (U04), followed by a content development project (PR01) originally also funded by the National Science Foundation – however, only few of these resources are in active use. Among the top resource providers are one public high school (HS20) and a content development collaboration of four public high schools, two middle schools, and one university (PR06).

The top six resource user institutions are all universities, where U05 is another large state university. Eight of the top twelve resource users are also among the top resource providers, but only U01 and U04 are making more resources available to the pool than they are using. All institutions in the top twelve with the exception of U03 (a medium-size public non-US institution) are using more resources from their own institution than from any other single institution, but only for U01, U04, and HS20 the absolute majority of used resources is generated locally.

OUTLOOK

This year, the LON-CAPA Academic Consortium was founded with Michigan State University and the University of Illinois at Urbana-Champaign as founding partners. Through this consortium, stewardship of the resource pool and the system infrastructure will be shared across institutions, and the financial sustainability of the project is guaranteed through long-term financial commitments of the partner institutions.

CONCLUSIONS

Resource sharing is a reality and can be accomplished across institutional boundaries. However, it takes a critical number of faculty (in this case around the initial randomizing homework project), a critical mass of resources (in this case, again, first around homework problems and built up over 14 years), and a long-term committed institution (in this case, Michigan State University) to get it started.

Resource sharing can indeed be a give and take, where the majority of the institutions making the strongest use of the system are also found among the strongest providers of content. While the driving force of our system are the online problems, the type of resource which on the average is used at the most institutions are animations and simulations, probably due to the high cost of development and their general appeal.

ACKNOWLEDGMENT

We thank the National Science Foundation (ITR-0085921) for the support of this project. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Support in earlier years was also received from the Alfred P Sloan and from the Andrew W Mellon Foundations. We are grateful to our own institution, Michigan State University and to its administrators for over a decade of encouragement and support, as well as our partner institutions.

REFERENCES

[1] E. Kashy, S. J. Gaff, N. Pawley, W. L. Stretch, S. Wolfe, D. J. Morrissey, and Y. Tsai, "Conceptual Questions in Computer-Assisted Assignments," *Am. J. Phys*, **63**(11), 1000 (1995)

[2] G. Kortemeyer and W. Bauer, "Multimedia Collaborative Content Creation (mc^3) - The MSU Lecture*Online* System", *J. Eng. Educ.* **88** (4), 405 (1999)

[3] G. Kortemeyer, W. Bauer, D. A. Kashy, E. Kashy, and C. Speier, C, "The Learning*Online* Network with CAPA Initiative", Proc. IEEE Frontiers in Education **31** (2001), see also http://www.lon-capa.org/