Open Sharing at MIT
Hal Abelson

To cite this version:

HAL Id: edutice-00001432
https://edutice.archives-ouvertes.fr/edutice-00001432
Submitted on 10 Mar 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
I am going to talk this morning about two programs that we have at the MIT that are based on the idea that part of the mission of the University is to share its knowledge openly with the world and to realize the power of the Internet for open sharing and collaboration.

There have been a lot of questions, a lot of discussions today and also five years ago, about what Universities should do when faced with the Internet. We asked that question at MIT also in 1999. When reporting on the deliberations that we did in the MIT, in answer to the question ‘What is MIT going to do about the Internet?’, the former president of the MIT, Mr. Charles Vest, said that the answer that the professors at MIT arrived at is that we are going to use the Internet to take all our course material and put it on the Internet for free for the entire world and forever.

That was a project that we announced in 1999 called “OpenCourseWare (OCW)”. We worked on it for more than a year before we announced it publicly, and the idea is what I said, MIT will take all of its courses, graduate and under-graduate, we put them on a Web site, the Web-Site would be public for the entire world, and this will be a permanent commitment from MIT. So part of what the university does is pay professors, part of what it does is pay for electricity for the rooms and part of what MIT would do is maintain permanently, distributing our course material for the entire world. MIT currently has about 1800 courses, and there are now about 1200 courses on the MIT Open Courseware site.. For every course in MIT, we have lecture notes, we have the exams, we have answers to the exams.

When we first put up OCW we would get e-mails from faculties around the world saying thank you very much for putting up the exams on the Internet, could you please take down the answers? And we said no, this is not for the world to use MIT’s material only, this is for
people to create their own material. We have some examples of text-books, some professors have elected to put complete text-books on OCW, for free; we have some video courses, a very small number, perhaps a dozen video-courses, but these are very expensive for us to produce and distribute. So many of the people would like us to put video courses on there, but again, remember, this is meant for MIT to pay for putting all of its courses on the network for free, so we could not support that on our budget, but we do a couple of special courses.

Putting your course material on OCW as a professor is completely voluntary. One of the things when we were designing OCW, we didn’t know how the faculty would react, so we said, to put your material on OCW is completely voluntary, and we felt that if it was successful, there would be social pressure for more and more people to put their courses on, and at the moment more than 70% of MIT faculty members have courses on OCW.

A second thing that we were very concerned about was issues about “Was the university going to appropriate faculty material?”. Because people would be very afraid that the university would take away faculty material and publish it. So on the OCW, the individual professors retain copyright on their material.

OCW was published under a licence, provided by an organisation called “Creative Commons” that I’ll speak about later, but the terms of the licence are that anyone may copy this material for non-commercial use. But also anyone may change it, and make derivative works. Because again the point of this is not that everyone should use the MIT material. The purpose of this is that people should be able to build on the MIT material and make even better stuff.

Let me say a little bit about how we started the project. MIT has a group called “The council on educational technology” that I direct. The council was formed in 1999 when there were questions about “oh, there is the Internet” and what should MIT do, what should we do about e-learning, what should we do about distance education. So we started an engagement with an American consulting company called McKinsey, just to look at the entire area, and we came up with a lot of different strategies. One was to increase the act of learning and MIT to reduce the number of large lectures, and we’ve done that. But also there was a recommendation to make stronger ties with MIT alumni. And maybe we said we can do that by developing courses. That would be offered over the Internet for MIT alumni, we called that “life-long learning”.

And then we started another engagement with another consulting company called “Booz-Allen-Hamilton”, and we worked the entire summer, trying to make a business model, for whether MIT could successfully do life-long learning. And the conclusion of that study which was very detailed and very quantitative, was that pretty much MIT would just lose money doing this. That unless you can imagine making a course for 25 000 students, you would probably lose money on it. And in any case it would take 5 years, 2 million dollars of up-front-investment and 5 years before the plan would even break even. That’s just a little example of the kind of analysis that we did, you see the graph of the number of corporations that we might work with, the number of professional organisations we might work with, and you see mostly for any of those assumptions about numbers of things and … you end up losing money. So, it came to the end of the summer and the study group was sitting there having discussed all these business plans, on the one hand you don’t make any money, on the other hand by starting to commercialise your course material, you run into all kinds of problems. So for example if MIT were to commercialise its course material, would we allow our students to share their courses with students at other universities? This is a big problem. So we came up with a different idea. We said since we are not going to make
any money on this, why don’t we just give it away for free? And what you see here is the actual recommendation:

“Conclusion 1: “This area is complex, highly competitive, rapidly evolving, and often not profitable.”

- An operation would lose money for MIT unless it was large (25K participants, 25 corporations, 5K alumni)
- Would require an investment of $2M and 5 years to break even.”

And also about the LLL Study, Executive Summary:

“VII. A revolutionary notion of OpenCourseware@MIT could radically alter the entire LLL and distance learning field and MIT’s role in it and should be seriously considered.”

From presentation to MIT administration, Oct. 2000

That recommendation was made to MIT’s academic council which is the governing body for the university. And the idea was, again this is the original idea, and again the idea was that we put up, we put up all our courses, it would take maybe 5 to 10 years to put up all these courses, we will try to fund it from foundations or government, and then we said along with that, maybe if we do a lot of free material maybe we’ll put up some more life long learning courses for pay, at the end of the day we never did do the life long learning courses for pay, because we still don’t think that there is a business model there for MIT, but open courseware expanded faster than we thought it would.

The MIT administration was really enthusiastic about the OpenCourseWare idea, but they said you cannot do this without going to the individual departments to see what they say. So
there was a group of about 5 of us who went around and gave presentations to MIT’s 33 departments. And we were very nervous of course as to what faculty might say, but overall it was very supportive. What I’ve listed here were the negative reactions.

**Negative reactions:**
- OCW deprives MIT of large potential monetary benefits (very few)
- OCW is not about education/pedagogical innovation, an area MIT should take leadership
- The OCW implementation and estimates of resources needed etc are not (well) developed and may not reflect (MIT) realities
- What will I do then in the classroom and will students still come to class?

**Positive reactions:**
- OCW will help many to do (basic) content modernization, providing resources and unburdening faculty from technical production chores.
- We are doing it anyway, might as well get some resources/support. Doing it on a large scale may be more efficient (technically, policy etc)
- (Few) OCW opens a new way for us to do knowledge dissemination and intellectual interactions: “we can’t afford not to do it”.
- (Few) OCW is a way for MIT to take the moral high road and exert leadership.

The negative reactions were few, very very few people said “by giving this material away you’re depriving MIT of income”. That’s a very small number, perhaps 10 members of the faculty thought that. Other people said this is too modest, you should do something much more technically ambitious, and if you do these posting lecture notes for all the courses you’ll be diverting energy from doing something better technically. Other people didn’t believe the business estimates we made, and other people said something very funny, they said if I take all of my lecture notes and post them on the Web, then why should students come to class? And the answer of course, and I hope as researchers and e-learning you give this answer too, is that if a student can get out of your class what they could get by looking at something on the Web, why are they going to class now?

So OCW really was partly an assertion by MIT that an MIT education is not this electronic material. The core of an MIT education is the interaction that students have with the instructors, and students have with each other. And if you consider that this stuff that we’re going to put on the Web, all these electronic things, that’s merely content. And our attitude towards that is that it’s just free. And it’s a real statement about the value of interaction between students and faculty, and a value of education or community in education. As I said, most of the reactions from the faculty were positive, and so another motivation was “if MIT does this big initiative of posting things on the Web”, and remember, this would be for every course of MIT, another motivation was “if we start doing the infrastructure for that, that would encourage modernisation across the whole institute of MIT”. Some of the faculty talked about open sharing, like I will, but that was a minority. Some of the faculty talked about MIT leadership.

But we went ahead, we decided to go ahead in March, in April we announced it. And we got a very very positive reaction. There was a large article on the front page of the *New-York Times* “MIT giving away all of its course material for free”. We had reactions from all around the world about what a wonderful altruistic thing it is. But, the point I want to make in the talk this morning is, this is not something that MIT did for altruistic reasons. We did this to have an impact on the climate that is causing educational materials, and academic discourse all together, to become to be commercialised. I’ll talk about that in a minute.
But first let me say a little bit about the impact of OCW. Here is data from last month. We have over 500,000 visitors per month.

The visitors, this means someone uses, someone visits OCW from some computer and then there is at least a half an hour gap between another visit, and by the way, the average session lasts about 11 minutes. Typically someone will look at 9 pages on the course Web Site. The average visitor visits the site about one and a half time a month. We distribute OCW through the Akamai network and service all around the world. Here is some data from Akamai, about who accesses in the MIT Site.
The numbers are simply web hits, so that doesn’t mean a whole lot, but you might look at the percentages from different places around the world, we have users from literally every continent included several users in Antarctica. Almost every country in the world, about half from North-America, a lot from China, 13% from Western Europe, and then goes down from there.

So this is a lot of users. What the people use it for? About half of OCW users describe themselves as self-learners. They are coming to the Site just for their own enrichment, and trying to enhance their personal knowledge. About a third are students from other universities who are looking for supplementary material. They are already taking courses and they are looking for some kind of supplement from the MIT courses. And 13% are educators, and that’s a large number, if you consider how many more students there are than educators.

So 13% of our users are educators and they use it for planning new courses. And in fact when we proposed OCW, this was the audience we were mostly interested in. This was meant to be MIT having an influence on other educators, in giving them resources for what they teach. We actually were a little surprised that they are so many self-learners. Because originally our course was not designed for that. We have lots and lots of awards, and of course that makes us very proud, but what’s more important to us is not what MIT is doing, but what other universities are doing.

So what I’m showing you here, this is a Web-Site by an organisation called “Universia” (www.universia.net) which is a consortium of 100 universities in Spain and Portugal and in America, and they are translating the MIT courses into Spanish and Portuguese. This is a site from another consortium, of about 50 universities in mainland China, which is translating the MIT courses into Mandarin. There is another project like this, that’s based in Taiwan, which
is translating the MIT courses into classical Chinese, and if you combine the visits to MIT, Universia, the Chinese Site, that’s called “Core” ([http://www.core.org.cn/cn/](http://www.core.org.cn/cn/)) and the Taiwan Site, there are about a million visits a month, of people coming from around the world looking at the MIT material.

Even better than that, the universities that are now publishing their own material for free, so the top eight Japanese universities are placing their own course material, this is Keio university ([http://www.keio.ac.jp/](http://www.keio.ac.jp/)), which is publishing courses from Keio, both in Japanese and in English. So all eight Japanese universities are doing that. If you look around the world, there are about 40 institutions now that are publishing their own material on the OCW Web Site. These are not MIT material. These are other places that have put up their own Web Sites for open sharing.

I must also mention ParisTech OpenCourseWare project. ParisTech is a collective entity that includes 11 of the most prestigious engineering schools in France. The research and teaching at the 11 institutions is complementary and cover virtually all sectors of the various engineering sciences ([http://paristech.polytechnique.fr/catalogue/?langue=EN](http://paristech.polytechnique.fr/catalogue/?langue=EN))

This is a graph of how many courses are available now over the last couple of years, and the nice thing is that there is now starting to be a sizable number of courses that are not MIT. And of course we are looking forward to an environment where MIT courses are just a small minority of all the free courses available on the Web.
Let me switch now and talk about another project at MIT. This one is called “DSpace” and it’s not only about course material, it’s about the university supporting an archive, a repository, where members of the university education and research community can publish their material. Because right now, if you write a paper, as a researcher, and you don’t give it to a publisher, but rather you place it on a Web server, for example, under your desk, but then there is no place where there can be a permanent commitment to making your work part of the academic record. Our vision for DSpace is that the university libraries would provide a permanent place, where the education and research contributions of the institution can be published.

The DSpace site is divided into several so-called communities. Communities are represented by different parts of the university. So we can go and look at one community, this is the community maintained by MIT’s business school, the Sloan school of management. Each community has several collections. So we can look at some collection, and see the offers, we’ve publications there, and we can pick one of them. And then we get a reference page, so this is some bibliographic information about the selection we just picked.

And there is the important part, that is what’s called a handle. A handle (e.g. http://hdl.handle.net/1721.1/1857 ), as some of you know, is like a web URL but it’s designed to be permanent. So if the machine changes, or the machine-address changes, that handle is still the valid reference, and what that means is that this paper can now be cited as part of the permanent academic record. So this is a permanent academic citation to this piece of work, and the vision is that this kind of extension, of the academic literature, would be maintained.
by university libraries. And of course the entire paper is in the repository, not only the reference. DSpace was a project carried out jointly by the MIT libraries, by the World Wide Web consortium, and by the Hewlett Packard company, which is interested in publishing. The technology itself is an open source content management system, it uses a standard called “RDF” for meta data, that’s being promoted by the Web consortium. RDF is a standard for inter-operable metadata. And like most open source software the DSpace software has been downloaded a lot of times. So since we put it up a couple of years ago, there have been about 6000 downloads.

What goes in DSpace? Well, all sort of things go in there. We initially designed it for research articles like the one I showed you, but now people are putting conference papers and videos and essays and e-learning objects and all sort of things. And also DSpace will be the permanent archive for OCW. Because if you go to the OCW Web site what you’ll find are MIT’s current courses. But if you’d like 10 years from now to find what MIT had been teaching in 2005, that would be in the permanent DSpace archive, maintained as a formal part of the MIT library collection.

I said there was 6000 downloads, that doesn’t mean a whole lot for open source, but what’s more important is that people are making their own DSpace implementations. So at the moment there are over 100 organisations that are putting up DSpace instances or repositories. I just show you, I’ll just flash through some slides that show a couple of them: from the Netherlands, Hong-Kong, Montreal, Ray university in India, from Milan, Parma, DSpace in Cambridge, and I can go on. As I said there, are over a 100 of these. And the result is that you have an inter-operable set of repositories that make the research contributions of the world’s research institutions available to everyone in the world.

So let me switch now and say a little bit about why MIT is doing this. These are two projects from MIT and I said they are not about altruism. They are about, I think, defending academic values. I think that if universities do not do these sorts of things that academic values are going to become increasingly marginalised. And people are going to start talking about the university as if it was nothing more than a business. Let me give you an example, you may remember about 5 years ago, when Napster became popular, all sorts of people were using Napster to share musical recording files, you might remember that. And a couple of musicians, Metallica and Doctor Dre, started sending out letters of complaint to universities, saying students are using your university networks to steal our music and such and such and such. Here is one reaction; this is a letter in response that the Dean of library at the university of Southern California sent to all students at USC. And to me this is a remarkable statement.

As an academic institution, USC’s purpose is to promote and foster the creation of intellectual property. It is antithetical to this purpose for USC to play any part, even inadvertently, in the violation of the intellectual property rights of others.

September 2002, letter to USC students from the Dean of Libraries

It says that the purpose of the university of Southern California is to promote the creation of intellectual property. Now that’s an amazing statement because I would have thought that the purpose of the university of Southern California had something to do with education. But apparently not. Apparently people become so enamoured, wrapped up in this notion of content and intellectual property that we forget what the university is actually about. To the point where a top administrator of university can make statements like this outrageous
statement. Here is a quote from Bill Bowen, who is just retiring now as president to the Mellon foundation.

“...one of my greatest concerns is that, either inadvertently or by design, universities will be so bemused by market opportunities that they will lose sight of, or downplay, their most essential purposes...”

William G. Bowen
President, Mellon Foundation
At a Slight Angle to The Universe Romanes Lecture, 2000

That’s talking about university how is getting too enamoured of this sort of e-business, e-learning business opportunities, to the point where they really forget what they are essentially about. This is essentially about academic communities and bringing a new generation into the culture of the world, it’s not about stuffing information into people.

Let me show you another example, which is my favourite. What you’re seeing here is a memo from the general counsel at the university of Texas in Austin, who is writing to the professors at the university of Texas saying: some of you have a problem, students sit in your class and they take notes in your lectures, but you must understand that if when they take these notes they don’t restrict themselves just to the fact, if those notes might contain some of your original expression, then your students are engaged in copyrighting infringement.

“Many students probably create a work that would infringe a faculty member's copyright, that is, they base their notes on and incorporate her particular expression rather than just state facts and ideas she articulates in more detail. Faculty members have always permitted this kind of activity without actually talking about it. They “implicitly” license students to create a “derivative work” from the lecture. The license is implied through academic tradition -- students are expected to take notes. ...

Now faculty may wish to make the implied license explicit and add some restrictions.

A limited license to take notes could be very important to protecting the intellectual content of lecture materials ...

University of Texas, Office of the General Counsel, August 2001
http://www.utsystem.edu/ogc/intellectualproperty/lectures.htm

Now, the memo goes on, how can we deal with this? Because on the one hand as professors we’d like our students to take notes in our lectures, but how can we condone copyrighting infringement? What do we do? So being a lawyer, we all know what lawyers do, the lawyers suggest you can solve this by having students agree to a license when they start the semester. And here is the suggested licence.

“Written and verbal instructions at the beginning of class could look something like this:
My lectures are protected by state common law and federal copyright law. They are my own original expression and I record them at the same time that I deliver them in order to secure protection. Whereas you are authorized to take notes in class thereby creating a derivative work from my
lecture, the authorization extends only to making one set of notes for your own personal use and no other use. You are not authorized to record my lectures, to provide your notes to anyone else or to make any commercial use of them without express prior permission from me.

University of Texas, Office of the General Counsel, August 2001
http://www.utsystem.edu/ogc/intellectualproperty/lectures.htm

So when students are walking to your class you inform them that they have the right to take notes for strictly personal use: you’re not allowed to share them with anyone. You know, you start wondering what the university is about, I mean maybe you say you’re allowed to sit in class and listen, but you’re actually not allowed to use that information, or something. It has all gotten a little bit crazy. And although this example is maybe a little amusing, it’s very serious. There is an excellent book about this by Corynne McSherry, who wrote this in 2001 when she was at Stanford Law School called “Who owns academic work?”. The key idea of the book is that there really is a difference between the way that our scholarship uses these words like “author” and “creator” and “originality” and the way that the copyright law thinks about authors and creators and originality, and the problem is as academics we become too enmeshed in the language of intellectual property.

Conflating “freedom of inquiry” with “freedom of property”

“Intellectual property law ... embodies the notion that the only forms of cultural work that can be “protected” are those that can be owned. ... ... the conflation of property rights and “academic rights” participates in a set of discourses ... in which freedom can only be understood to mean “individual free enterprise.”

In retelling this tale academics risk losing a language for talking about knowledge as other than private property and the university as other than economically “useful.”

Corynne McSherry, Who Owns Academic Work? (2001)

When we talk in legal terms what it is that we do in classes, and the net result of this is that the university is going to lose something very special. And the more we go down this route, the more eventually the university will be seen as just another almost commercial institution and lose its special place in society. So what we need to create, and this is what OCW and DSpace are about, is some world where there are open educational resources, some world where there is high quality content that allows access to knowledge.

There are many many problems in doing this, some of them has to do with interoperability and removing barriers. So there is ideas of, if we put up content, how do we translate it? How do we localise it? Can you imagine content put up with tools that do automatic translation, of course material? There are intellectual property problems. How do we arrange the intellectual property licensing so that people across different countries can really build on each others work?

What you need is infrastructure. Some infrastructures are technical, the things like a DSpace content repository, there are things like an OCW publishing system. But there’s also legal infrastructure that is required.

I mentioned before that if you go to MIT OCW you see a Creative Commons licence. This is an example of what a licence looks like. Let me show you the version from Creative Commons France, this is a particular kind of licence. Creative Commons is a family of licences that allow different authors to make different choices. This particular licence says
you may copy the work and redistribute it, but you must preserve attribution, you must say who the original author was. You may not use it for commercial use, and if people make derivative works they have to distribute those derivative works under the same licensing terms as they got it from. We call that “share alike”. But there is a whole collection of such licences. How do you attach such a license to your work?, You go to the Creative Commons Web Site, there are Creative Commons Web Sites around the world, take the French one (http://fr.creativecommons.org/), and you answer some questions about your work: do you want to allow commercial use? Do you want to permit people to make derivatives? And so on. And what you get back is an image that you put on your web page, a logo with the words “some rights reserved”.

That image is linked to a version of the licence that we call “the Commons Deed”, this is a very simplified form of the licence which is meant to be understandable by people without legal training. That Commons Deed is linked to an actual formal licence which is designed to be a very proper copyright licence in harmony with the copyright laws of different countries. And that licence is linked to machine readable code in a language called RDF developed by the Web Consortium, and it says such things as: this particular page permits people to make derivative works or requires if you redistribute it attribution to the author.

The consequence is that when you put machine readable code on there, people can start building infrastructure to take advantage of the code. At the moment there are about 60 million documents on the Web that have Creative Commons licences. There are tools, for example an extension to a Mozilla Firefox browser, that looks at the Web Page and signals by these little icons, what its licence status is. You can also build search engines. So you can say, don’t just go on the Internet and do a search for something. Do research for material that’s free, or do a search for material that I can modify. One of those is Yahoo, so if you go to Yahoo Search or Google search, either one, you can say search for material that’s license standard Creative Commons licence. So for example I can go to the Yahoo search page and say “find photographs of a gargoyle at Notre-Dame”. Find content that I can copy and modify. I go to the search page and back comes a very very bad photograph of the gargoyle of Notre-Dame. But the point is, even though it’s a very bad photograph, you’re free to use it and redistribute it and put it in your work and put it in your books. That particular photograph is on Flickr (http://www.flickr.com/) which is a photograph sharing site on the Web, and Flickr lists things under different Creative Commons licences and there are about 5 million Creative Commons licensed photographs that you can use for free, as part of Flickr.

We’re also in Creative Commons starting a project called “Science Commons”. “Science Commons” takes the same ideas of sharing and applies them to academic publishing and university licensing. We’re doing a lot of work in genetic databases, the terms under which biologists can share genetic material.

One of the difficult approaches in copyright is harmonising copyright across different countries. So there are Creative Commons licences in about 20 countries so far and projects for more. And the notion, the vision is that there is a world wide pool of content that not only for people to use but for people to reuse and build upon. We envision that you unlock the
power of the Internet for joined collaboration and sharing, and in such a way that really supports academic values.

So let me conclude and say universities have real core institutional reasons to support open education resources. It’s not just a nice altruistic thing to do, it really is about defending academic values. Not only that, but universities really can establish infrastructure. A university can make an open courseware Web Site, a university can make a DSpace, a university can put out its material so that the rest of the academic community can build upon. And the third point is that everyone can support these open resources by using Creative Commons licences.

Thank you.