May 16, 2009

How to Find, Use, and Get Open Access Content

Laura Quilter, *University of Massachusetts - Amherst*

How to Find, Use, and Get

Open Access Content

Laura Quilter / Simmons College

2009 May 16

Roadmap

1. Surveying the landscape
    - concepts, terms, licenses
    - FAQs
    - *discussion questions / worksheet*
2. Finding open access content
    - databases, googling, keeping up
    - grokking the meta-data
    - *discussion questions / worksheet*
3. Using open access content
    - technical how-to
    - of particular interest to libraries
    - *discussion questions / worksheet*

Wrap-up & Q&A

Surveying the landscape

part I – answers to these questions:

- *What's the difference between open access, open source, public domain, copyright, copyleft, and Creative Commons?*
- *What can we do with different types of licenses?*

"On the one hand **information wants to be expensive**, because it's so valuable. The right information in the right place just changes your life. On the other hand, **information wants to be free**, because **the cost of getting it out is getting lower and lower all the time.** So you have these two fighting against each other."

Stewart Brand
first Hackers' conference, 1984

Surveying the landscape

1. concepts & terms

Information may want to be free, but people can make a profit from closing it up & charging for it. What's not "open"?

rules

- copyright
- trademark
- patent
- "hot news"
- databases
- govt "classified"
- secrets: trade secrets, NDAs

technology

- private, unpublished
- obscurity
- "technological protection measures"
- encryption

Surveying the landscape

concepts & terms

- copyright :
  - 6 exclusive rights accorded to copyright holders; copyright *automatically* attaches when a work of original authorship is fixed in any tangible medium. [–]
    - reproduction
    - distribution (first distribution)
    - derivative works
    - perform
    - display
    - certain digital transmissions
  - Copyright terms extended over repeatedly; currently, life+70.
  - major exclusions: [+]

Surveying the landscape

concepts & terms

- copyright :
  - 6 exclusive rights accorded to copyright holders; copyright *automatically* attaches when a work of original authorship is fixed in any tangible medium. [+]
  - Copyright terms extended over repeatedly; currently, life+70.
  - primary limitations: [–]
    - fair use!
    - first sale!
    - public domain!
    - idea-expression dichotomy!

Surveying the landscape

concepts & terms

primary limitations on copyright:

- fair use! *protects many uses for:*
  - educational uses
  - parodic & critical uses
  - transformative uses
  - limited uses
  - reverse engineering
- first sale! *protects your right to re-sell or lend your copy*
- public domain! *pre-1923, expired, government works*
- idea-expression dichotomy! *facts, information, generic plots*
- Section 108! *library exceptions*

Surveying the landscape

copyright and licensing

Copyright establishes the default allocation of rights between copyright holders and users of copyrighted works.

But copyright holders use *licenses* to vary the defaults.

Licenses = permission to use, under conditions.

* usually used to restrict the user's rights under copyright: no reverse engineering, no copying, no quoting, etc.

Open Access Licenses = permission to use, under conditions.

* inverts the typical paradigm to grant *more* rights to user
* clever idea; widely used; subject to some controversy

Surveying the landscape

common licensing terms

restrictions

- no reproduction
  - *not even for personal*
  - *no reverse eng.*
- no distribution
- no derivative works
  no distribution of DWs
- particular to the medium, e.g., video game alterations; software use restrictions

"open" response

- free to copy
  - *reverse engineering*
- free to share
- free to "remix", tinker, DWs
  - *free to distribute DWs non-commercially*
    - *commercially*
  - *required to distribute*
- particular to the medium

Surveying the landscape

open response in software

Richard Stallman's Four Freedoms of "free software":*

(0) freedom to run the program;
(1) freedom to study & adapt the program (i.e., source code);
(2) freedom to redistribute;
(3) freedom to improve & distribute

* Not the FDR version. "Free Software Definition", Richard Stallman, Free Software Foundation (1986), http://www.gnu.org/philosophy/free-sw.html

Surveying the landscape

open response in software

free software vs. open (source) software

- "free software" : software is free; services, documentation, copies/shipping, *value-added* can cost $$. "free as in freedom, not free as in beer"
- open source specifies *access to the source code* and *ability to distribute* and *ability to modify*
- people who don't care or want to talk about it all: *"FOSS"*

Surveying the landscape

concepts & terms

- free software - *software free to read, use, modify, but you have to distribute modifications as free software*
- open source sw - *software free read, use, and modify*
- "shared source" - *look but don't touch*

- open access - *free for all readers to read & use*
- open content - *free to access, manipulate, distribute*
- open publishing - *free to all to contribute; e.g., Wikipedia, IMC*
- open standards - *free to implement; openly negotiated*

Surveying the landscape

concepts & terms

- copyrighted - not in the public domain; subject to licensing
- licensed - permission to use under certain conditions
- *copyleft* - inverts the licensing paradigm to grant more rights to user

- public domain : free of legal restrictions, esp. in copyright; e.g., copyright term has expired
  - "genericide" (trademark)
  - "out of patent"

- "the commons" : the set of material that is in the public domain

Surveying the landscape

concepts & terms

open access =

- copyleft & licensing on generous terms
- materials *in* the public domain
- materials *committed* to the public domain

Surveying the landscape

2. common licensing terms

additional rights

- copy
- share / distribution
- derivative works / remixes

additional "rights"

- attribution
- share-alike / viral clauses
- non-commercial use only

Surveying the landscape

2. common licensing terms

two major schools of open licensing:

- "copyleft" - includes the viral "share alike" clause
  - modifications must also be released copyleft
  - GPL, GFDL, Creative Commons SA
- "permissive licensing" - excludes viral clauses
  - permits incorporation into non-openly licensed material
  - BSD, MIT, sleepycat

Surveying the landscape

3. common licenses

Beyond the two major schools, a myriad of proliferating licenses

- "Free Software Foundation" defines free sw licenses
- "Open Source Initiative" defines open source sw licenses
- Creative Commons - established for texts, graphics, etc.

- dedicated to the public domain (*not* a license)

Surveying the landscape

3. common licenses

GPL: "GNU General Public License"

- originally written by Richard Stallman in 1989
- now on version 3 (with significant input from Software Freedom Law Center)
- "Four Freedoms"
- managed by Free Software Foundation
- requires transfer of ownership to FSF *for some projects*
- major players: Richard Stallman, Software Freedom Law Center (http://softwarefreedom.org/), Eben Moglen
- distinguishing feature: *software; copyleft/viral clause*
- see also GNU Lesser General Public License

FREE SOFTWARE FOUNDATION

Surveying the landscape

3. common licenses

GFDL: "GNU Free Documentation License"

- managed by Free Software Foundation
- intended for text documents, e.g., software documentation
- biggest user: Wikipedia
- compatible with Creative Commons
- current version 1.3
- anti-DRM clause
- distinguishing features: *text*-oriented, copyleft/viral clause

Surveying the landscape

3. common licenses

- BSD licenses: "Berkeley Software Distribution"
  - developed by University of California for BSD OS
  - a "permissive" license
  - simply requires inclusion of licensing information
    - original clause required "UC" on advertising
  - distinguishing features: the classic "permissive" license, with *very few restrictions on use*
    - e.g., can incorporate BSD software into commercial SW
- MIT / X11 license:
  - a "permissive" license
  - GPL-compatible
  - frequently "dual-licensed"

Surveying the landscape

3. common licenses

OSI licenses

- certified by Open Source Initiative
- includes both "free" (like GPL) and "permissive" (like BSD)
- also includes most public domain software (as long as source code is available)
- "open source" chosen advisedly as counter to activist term "free"

Surveying the landscape

3. common licenses

Creative Commons

- developed beginning 2001
- key people: Larry Lessig
- distinguishing features: *cultural & scholarly works*; mix-and-match licensing terms; free distribution/copying
- key licensing options:
  - attribution: "BY"
  - derivative works must be share-alike: "SA"
  - no derivative works made: "ND"
  - non-commercial uses only: "NC"
- see also Science Commons, Education Commons, etc.

Surveying the landscape

3. common licenses

resources to learn more

- *Open Source Licensing* by Lawrence Rosen
- *Understanding Open Source and Free Software Licensing* by Andrew St. Laurent

- Free Software Foundation: http://www.fsf.org/licensing/
  - free software definition: http://www.fsf.org/licensing/essays/free-sw.html
- Software Freedom Law Center: http://softwarefreedom.org/
- Open Source Initiative: http://opensource.org/
  - open source definition: http://www.opensource.org/docs/osd

Surveying the landscape

FAQs

- Are open access licenses really enforceable?
  - yes ... ? *Jacobsen v. Katzer* (Fed.Cir. Aug. 2008) - "Artistic License 1.0" was an enforceable copyright license as applied to infringement of Java Model Railroad Interface (JMRI)
- Are these licenses incompatible with each other?
  - Sometimes, yes. *Dual licensing* (or *multi-licensing*) can help.

Surveying the landscape

FAQs

- If I license my material open access, am I restricted in what I can do with it?
  - No; you can use *dual licensing* to license it openly *and* to sell it or license it in a different way. Open access licensing can co-exist with any other kind of licensing. You can give it away in one market and sell it in another.

- Can I sell someone else's openly licensed content?
  - No, you can't sell the content itself, but you can sell value-added services: documentation, service, nice copies...

Worksheet & Break

Finding open access content

part II – answers to these questions:

- *What are the key databases & archives for open access content?*
- *How can I use search engines to look for licensing meta-data?*

Finding open access content

databases, archives, repositories

- Internet Archive - everything, but see especially film & music. Includes public domain and openly licensed content.
  - http://archive.org/
- Creative Commons database - sound & video recordings, texts
  - http://creativecommons.org/
- iBiblio - wide variety of materials, arranged academically
  - http://ibiblio.org

- Open Access Directory @ Simmons - index of indexes
  - http://oad.simmons.edu/oadwiki

Finding open access content

databases, archives, repositories

- books: Project Gutenberg, Open Content Alliance, Google Books

- scholarly research:
  - Directory of Open Access Journals, http://www.doaj.org/ and JURN, http://jurn.org (arts and humanities)
    - e.g., *PLOS; First Monday*
  - self-archiving
    - Google Scholar; Science Commons
  - preprint/etc archives:
    - arXiv.org (first preprint service; physic); SSRN, BEpress, Project Perseus (humanities)
    - institutional/national: eprints.org (ROAR - Registry of Open Access Repositories)

Finding open access content

databases, archives, repositories

- biomedical:
  - *PLOS* / Mike Eisen
  - biomedical "delayed open access"
  - NIH Public Access Policy:
    - PubMed deposits w/in 1 yr of publication
    - Omnibus Appropriations Act of 2009 made it law
    - Conyers bill, H.R. 6845, "Fair Copyright in Research Works Act"
    - Alliance for Taxpayer Access

Finding open access content

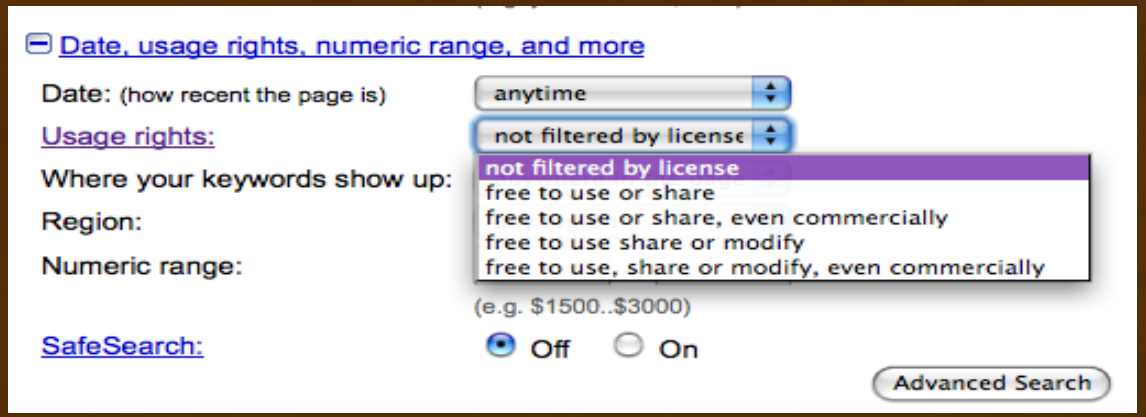
databases, archives, repositories

- text: scribd.com
- images: Internet Archive, CC database, government (e.g., NASA)
- music: Creative Commons (creativecommons.org), Internet Archive (archive.org), ccmixter.org
  - music data: MusicBrainz.org (open access replacement for CDDB, bought by GraceNote)
- video: Internet Archive's Moving Image Archive (http://prelinger.com/) , Open-Video.org
- software: SourceForge.net, FreshMeat.net

Finding open access content

usage rights in search engines

- Google Advanced Search > usage rights [searches for Creative Commons license data]

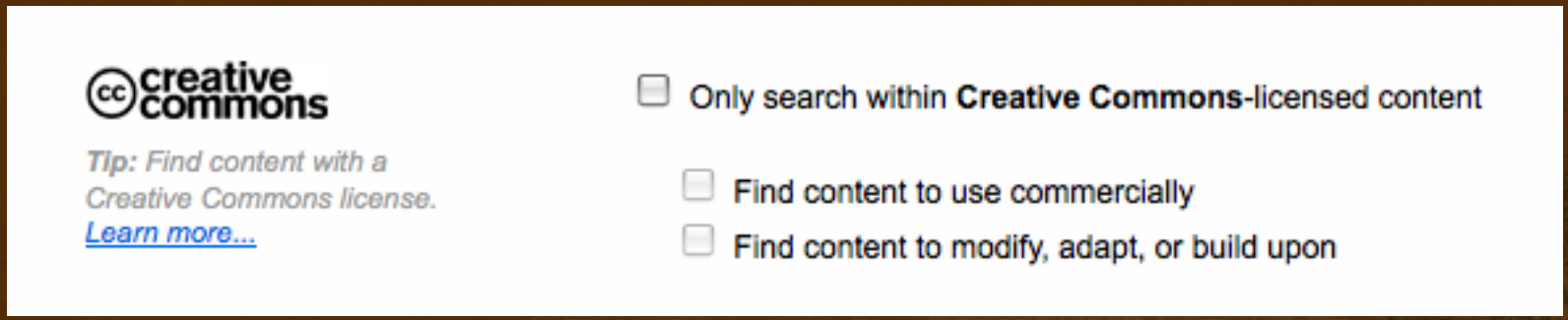


- Yahoo! Creative Commons Search http://search.yahoo.com/cc

Finding open access content

usage rights in search engines

- Flickr : The Commons http://www.flickr.com/commons/
- Flickr : Advanced Search



Finding open access content

keeping up

- Peter Suber, *Open Access News*
- SPARC
- Walt Crawford, *Cites & Incites*
- *Chronicle of Higher Education*, esp. "Wired Campus" blog
- *First Monday*
- *D-Lib*
- Simmons Open Access Database

Finding stuff

meta-data

- Rights Expression Languages specify standards for describing rights and license types.
  - XML-based: information described in a standard way that specifies particular attributes or properties
  - can be made to be "machine-actionable", i.e., DRM-ready

references:

- Karen Coyle, *Rights Expression Languages: A Report for the Library of Congress* (Feb. 2004), http://www.loc.gov/standards/Coylereport_final1single.pdf
- Gord Larose, *The DRM Dictionary*, http://www.info-mech.com/

Finding stuff

meta-data

- XRML: eXtensible Rights Markup Language - machine-actionable; sponsored by the content industry
- ODRL: Open Digital Rights Language - machine-actionable; sponsored by the W3C
- ccREL: Creative Commons Rights Expression Language
- METSrights: Metadata Encoding & Transmission Standards; Library of Congress/consortium
- ECMI: Electronic Resource Management Initiative
- Adobe Content Manager, Mobipocket (kindle), etc.

Finding stuff

meta-data

CCREL example:

My Book by Jon Phillips is licensed under a Creative Commons Attribution Non-Commercial 3.0 License. Permissions beyond the scope of this license may be available at <a xmlns:cc="http://creativecommons.org/ns#" rel="cc:morePermissions" href="http://somecompany.com/revenue_sharing_agreement">somecompany.com.

Worksheet & Break

Using open access content

part III – answers to these questions:

- *What are the step-by-step instructions for making something open access?*
- *What are five exciting open access initiatives for libraries?*

Using open access content

3 simple steps to creating open access content

1. figure out the terms & pick out the correct license
2. attach the licensing information in a manner appropriate to the format
3. distribute: publicize, archive, post

Using open access content

3 simple steps to creating open access content

2. attach the licensing information - *You can use one or more different ways of marking content with licensing data.*

- text: (1) Include a statement referencing the license and URL and/or the text of the license. (2) Electronic documents (PDFs) can also embed licensing information in the meta-data.
- HTML: (1) Include licensing data in visual portions of the website. (2) Embed licensing information in the meta-data.
- software: Embed licensing information in file headers, comment lines, and/or separate license files.

Using open access content

3 simple steps to creating open access content

2. attach the licensing information (cont'd)

- graphics: (1) Embed licensing information in the meta-data. (2) Add licensing logos or rights information to the image in a watermark, caption, or other visually accessible way.

- music: (1) Embed licensing information in the meta-data. (2) Add an audio "bumper" (a sound clip at beginning or end) describing the licensing. (3) Printed on physical media.

- video: (1) Embed licensing information in the meta-data. (2) Include a video bumper. (3) Print on physical media.

Using open access content

3 simple steps to creating open access content

2. EXAMPLES

- ccPublisher
- Adobe PDF
- browser plugins

more:
https://code.launchpad.net/projects/?text=metadata

Using open access content

3 simple steps to creating open access content

3. distribute: publicize, archive, post

- deposit in archives
- self-archive
  - Science Commons sample letter to publisher
- post online on sites that are indexed by search engines (e.g., a blurb in a blog or newsletter; update a list on a wiki; include appropriate cites in online reference material such as Wikipedia)

Using open access content

of particular interest to libraries

- bibliographic data:
  - Open Library - http:/openlibrary.org/ bibliographic data
  - Open Access Directory - bibliographic data
  - open classification project
  - OCLC records policy (Policy for Use & Transfer of ...Records)

- standards development:
  - OAI-PMH 2.0 (Open Archives Initiative Protocol for Metadata Harvesting)
  - RELs (see previously)

Using open access content

of particular interest to libraries

- library software:
  - Koha - open source OPAC
  - open source versions of major sw packages for internal productivity or public services: blogs (wordpress), CMS, wikis (mediawiki), personal bibliographic software packages, graphics (GIMP), MIT Open Courseware
- library catalog enhancements:
  - LibraryThing covers, APIs, Common Knowledge http://librarything.com/

Using open access content

ideas for using open access

- Celebrate Open Access Week (Oct. 19-23, 2009)
- Educate colleagues about open access resources & alternatives
- Educate patrons - community, faculty, students - about copyright & open access
- Participate in library-related open source projects (classification, OAD)
- Include open access materials in the library catalog, with links to the full-text, or locally archived copies of the full text.
- Set up a local or institutional "open access archive" that the library runs for your staff or members or community or faculty

Wrap-up

Me: lquilter@lquilter.net ...

this: teaching.lquilter.net/oa2009/

[userid/password: open/access]