Copyright & Digital Scholarship
Workshop for MLAtechcamp 2018
Portland, Oregon

Kathleen DeLaurenti
Head Librarian
Friedheim Library & Archives
The Peabody Institute of Johns Hopkins University
You know why you came....

YOU GET FAIR USE,
YOU GET FAIR USE

YOU ALL GET FAIR USE!
It’s not all about fair use
First things first

1. Do I have a copyright problem?
   a. Is the source material your working with *copyrightable*
   b. Is the source mater you’re working with *copyrighted*
Originality

Joseph Karl Stieler, Portrait of Beethoven, 1819, oil, Collection Walter Hinrichsen, NY.

The yellow pages
Is it fixed?

Ice sculpture by Ice Matter, LLC
Is it in the public domain?

WHAT IF I TOLD YOU

NOT ALL COPYRIGHT LASTS FOR 70 YEARS AFTER THE DEATH OF THE AUTHOR... nemegenerator.net
Where are we?

COPYRIGHT IS HARD

Is it copyrightable?

Yes

No

Is it copyrighted?

Yes

No

Is it fair use?

NO RIGHTS REQUIRED

DIGITAL SCHOLARSHIP SUCCESS
17 U.S.C. Sec 107

Notwithstanding the provisions of sections 17 U.S.C. § 106 and 17 U.S.C. § 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.
How do I know?

Over the years, courts have returned to two questions to decide if a use is fair:

1. Did the use “transform” the material taken from the copyrighted work by using it for a broadly beneficial purpose different from that of the original, or did it just repeat the work for the same intent and value as the original?

2. Was the material taken appropriate in kind and amount, considering the nature of the copyrighted work and of the use?

*Code of Best Practices in Fair Use for Academic and Research Libraries*
Examples of DH & Fair Use

HathiTrust

Google Books

www.hathitrust.org
More examples from the courts

- Quoting small amounts of text (8% or less) from each of 25 writings in a critical biography of L. Ron Hubbard was held a fair use in *New Era Publications Int'l ApS v. Carol Publishing Group* in 1990.

- Reproducing seven posters in significantly reduced sizes that had once advertised Grateful Dead concerts in a 480 page book on the cultural history of the band was held to a fair use in *Bill Graham Archives v. Dorling Kindersley* in 2006.

- Scanning student papers into a database designed to detect plagiarism was held fair use in *A.V. v. iParadigms* in 2009.

- Retelling the story of Gone With the Wind from the vantage point of a slave was held a fair use in *Suntrust Bank v. Houghton Mifflin Co*. (However, an unauthorized sequel to Catcher in the Rye, imagining Holden Caulfield as an old man, was held unfair in *Salinger v. Colting*. Sequels, in general, are likely to be considered infringing derivative works. One reason the use was fair in Suntrust was because of it was a critical commentary on the original and the Mitchell estate made clear it would never have licensed this kind of use of the famous novel.)
Digital Humanities Manifesto on Fair Use

Guerilla action item:

weak = ignore the well-intentioned “voices of reason” that will always argue for interpreting scholarly or artistic fair use in the most restrictive manner (so as to shield the institutions they represent from lawsuits, no matter how improbable or unfounded); adopt vigorous interpretations of fair use that affirm that, in the vast majority of cases, scholarship and art practice: a) are not-for-profit endeavors whose actual costs far exceed real or potential returns; and b) are endeavors that, rather than diminishing the value of IP or copyright, enhance their value.
More manifesto

-- the great diminishers: they will reduce anything in digital humanities (it's just a tool; it's just a repository; it's just pedagogy). They have rarely, if ever, built software, parsed code, created a database, or designed a user interface. They are uni-medium scholars (most likely of print) who have been lulled into centuries of somnolence.

--the false fellow travelers: they will wave the banners of change with continuity on their agenda. What's at stake is not simply continuity vs. change but honesty vs. hypocrisy.

--all those who would falsely equate the tools of the present with a turn away from history in the name of presentism, vougishness, or vocationalism

--the traffickers in IP

--university legal offices whose definitions of Fair Use amount to No Use

--archives, museums, libraries, and corporations that restrict access by means of cost barriers
Seek permissions OR check your license to see if it prohibits your use case.
Scenario #1

You want to produce a website related to some unknown children’s music by a Swedish composer. You have worked with some children to record the songs and would like to post those recordings, along with the poems and their translations, on your website. This will be accompanied by some history of the composer, these works, and some scholarly analysis of the pieces.

The music and poems are still under copyright; in fact, you happen to know that the composer never sought permission to set the poems. The poet wrote publicly about their frustration when the songs were published, but never pursued legal action. Both the poet and the composer are deceased.
Scenario 2

You are working on a project to create new chamber music arrangements of the works of underrepresented composers. As a scholar, you have traveled to archival collections, worked with archives to get access copies of original manuscripts and early editions, and based your arrangements on the most original version of these works available. Some of these pieces are in the public domain and some of them are protected by copyright. They were written between 1890 and 1960; not all of them were published. Your goal is to make a web resource of these arrangements available for free to encourage more diversity in programming and make it easier for performers to find works by underrepresented composers to perform.
Scenario 3

You are working with a computational musicologist. They are interested in analyzing musical datapoints like rhythm, energy, and timbre in conjunction with word frequency in lyrics for rap songs from the 1990s. There are two goals of the project: one is to publish the research in an open access publication, including the data sets analyzed in the project and the second is to create a Google Ngram type site that would allow other scholars and aficionados of 1990s rap to experiment on their own and create their own lyrics analysis based on the data set.
A tale of two data sets

musiXmatch dataset

The MXM dataset provides lyrics for many MSD tracks. The lyrics come in bag-of-words format: each track is described as the word-counts for a dictionary of the top 5,000 words across the set. Although copyright issues prevent us from distributing the full, original lyrics, we hope and believe that this format is for many purposes just as useful, and may be easier to use.

https://labrosa.ee.columbia.edu/millionsong/musixmatch

55000+ Song Lyrics
Lyrics for 55000+ songs in English from LyricsFreak

The dataset contains 4 columns:

1. Artist
2. Song Name
3. Link to a webpage with the song (for reference). This is to be concatenated with http://www.lyricsfreak.com to form a real URL.
4. Lyrics of the song, unmodified.

https://www.kaggle.com/mousehead/songlyrics
Resources


Reach Kathleen [@delaubrarian](mailto:@delaubrarian) or kdelaurenti@jhu.edu