What sort of copyright questions might you see?

• Researchers/authors using copyrighted content in articles, presentations, theses
• Sharing data and handling author’s own copyrights
• Text mining large corpora of copyrighted content (including library resources)
• What else?
What is copyright?

- Reproduction
- Modification
- Distribution
- Public Performance
- Public Display
What is covered by copyright?

“Original works of authorship fixed in any tangible medium of expression”

Including:

- **Literary works**
  - text
  - plot
  - software
  - characters

- **Music**
  - compositions
  - sound recordings

- **Visual art**
  - Paintings
  - Sculpture
  - photographs

- **Performances**
  - plays
  - choreography
  - motion pictures
  - audiovisual works
What ISN’T covered by copyright?

- Facts
- Ideas
- Insufficiently creative arrangement of facts or ideas
- Works in the public domain
Facts, ideas, (data), and the arrangement thereof

Image by Adam Lehman, CC-BY

Figures/data – example 1


<table>
<thead>
<tr>
<th>Race/Hispanic Response in 2010 Census Linked Data</th>
<th>Race/Hispanic Response in 2000 Census Linked Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>122,765,113</td>
<td>710,019</td>
</tr>
<tr>
<td>SOR</td>
<td>50,005</td>
<td>4,112</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>417,855</td>
<td>4,911,175</td>
</tr>
<tr>
<td>SOR</td>
<td>75,105</td>
<td>2,380,183</td>
</tr>
<tr>
<td>Any Other Response</td>
<td>848,876</td>
<td>135,580</td>
</tr>
<tr>
<td>Total</td>
<td>124,156,954</td>
<td>8,141,069</td>
</tr>
</tbody>
</table>
Fig. 2
Frequency of race and Hispanic-origin response change between 2000 and 2010. Census 2000 and 2010 census linked data; $N = 162$ million people. Response stability is indicated by □. Each cell that is shaded with values in the key above shows the number of people with that race/Hispanic response change. Quadrants 1 and 4 show race response change frequencies for people whose Hispanic/non-Hispanic response was stable. Quadrants 2 and 3 show those whose Hispanic/non-Hispanic response changed. Rows and columns are in census order (see Fig. 3 for partial list and U.S. Census Bureau (2007:6-1 to 6-3) for full list). Only cases in the linked data are shown. People reported as
Figures/data – example 3

Fig. 5 Race/Hispanic response change among people in the linked data, by sex, age, and race/ethnicity response in 2000. Charts include only people in the linked data. See Table S3 in Online Resource 1 for underlying numbers. AIAN = American Indian/Alaska Native
The public domain

- Works of the US federal government
- Works whose copyright has expired
- Works placed in the public domain by their creators
The public domain

Works whose copyright has expired

Copyright Term and the Public Domain in the United States

Never Published, Never Registered Works

<table>
<thead>
<tr>
<th>Type of Work</th>
<th>Copyright Term</th>
<th>In the public domain as of 1 January 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unpublished works</td>
<td>Life of the author + 70 years</td>
<td>Works from authors who died before 1948</td>
</tr>
<tr>
<td>Unpublished anonymous and pseudonymous works, and works made for hire (corporate authorship)</td>
<td>120 years from date of creation</td>
<td>Works created before 1898</td>
</tr>
<tr>
<td>Unpublished works when the death date of the author is not known</td>
<td>120 years from date of creation</td>
<td>Works created before 1898</td>
</tr>
</tbody>
</table>

Works Registered or First Published in the U.S.

<table>
<thead>
<tr>
<th>Date of Publication</th>
<th>Conditions</th>
<th>Copyright Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before 1923</td>
<td>None</td>
<td>None. In the public domain due to copyright expiration</td>
</tr>
</tbody>
</table>

https://copyright.cornell.edu/publicdomain
The public domain

Works placed in the public domain by their creators

https://creativecommons.org/share-your-work/public-domain/cc0/
Open licenses

ODC Open Database License (ODbL)
http://opendatacommons.org/licenses/odbl/
Attribution, Share alike & keep open

ODC Attribution
http://opendatacommons.org/licenses/by/
Attribution

ODC Public Domain Dedication
http://opendatacommons.org/licenses/pddl/summary/
None
Finding openly licensed content

Images

- https://search.creativecommons.org
- https://meta.wikimedia.org/wiki/Free_image_resources
- Google search filter by usage rights (and verify!)
So, the work is covered by copyright, and isn’t openly licensed, what do you need to do to use it?

**Use a copyright exception**
- Fair use
- Other exceptions

**Get permission**
- Directly from the copyright holder
- Via a collective licensing organization
The fair use of a copyrighted work [...] for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.

Fair use – 17 USC §107

Purpose and character of your use

- Transformative
- Educational research
- Scholarship

Nature of their work

Amount taken

Market effect of your use

Four factors
Fair use – example 1

http://rebekahmodrak.com/work/re-made-best-made-echo/
Fair use – example 2

A.V. et al. v. iParadigms, LLC, 562 F.3d 630 (4th Cir. 2009)

Reduce Plagiarism

Check students' work for improper citation or potential plagiarism by comparing it against the world's largest academic database.

Demo

A Natural Setting

A History of Exploration and Settlement in Yosemite Valley

Yosemite National Park is a United States National Park spanning eastern
A graduate student is researching representations of gender in YA literature over time. They’ve identified 25 novels from your library’s fiction collection that they would like to use for their analysis, ranging from highly popular works to some more niche, hard-to-find items. They want to scan the full text of the novels and collect data on every instance of third-person pronoun use.
The fair use of a copyrighted work [...] for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.

Four factors:

1. Purpose and character of your use
2. Nature of their work
3. Amount taken
4. Market effect of your use

Transformative, scholarship, educational research.
Fair use - exercise

Group 1 Question

Based on your analysis of the four fair use factors, would the scanning and data collection proposed by the student be a fair use?
The fair use of a copyrighted work [...] for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.
The student has collected the data, and is now in the process of writing up their research for their thesis. They would like to include their raw data as an appendix. The raw data includes every pronoun found in the books, and the 20 words immediately preceding and following each pronoun, to show the context of the use. Based on your assessment of the four fair use factors, would including this appendix in their thesis be a fair use?
The fair use of a copyrighted work [...] for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.

**Four factors**

1. **Purpose and character of your use**
   - transformative
   - educational research
   - scholarship

2. **Nature of their work**

3. **Amount taken**

4. **Market effect of your use**
Fair use - exercise

Group 3 Question

The student has collected the data, and is now in the process of writing up their research for their thesis. They would like to include the following cover image from one of the selected novels in their thesis. Based on your assessment of the four fair use factors, would this be a fair use?
The fair use of a copyrighted work [...] for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.
Other copyright exceptions

§110(1) – the classroom exception
- Performance or display in the course of nonprofit face-to-face teaching
- Applies to images, video, music

§108 – library exceptions
- Interlibrary loan
- Preservation copies

§512 – DMCA safe harbor for online service providers
- Applies when you provide a platform for unmediated, user-uploaded content
- Various requirements, notice and takedown procedure
Finally, what happens if someone gets upset, and what exactly is at stake?

• Monetary damages
  – Profits from the use
  – Statutory damages: $200 - $150,000
    • BUT! Not applicable against a school or library acting in a good faith belief of fair use

• Injunctions (i.e. stop doing that)

• Most cases will start with a cease and desist letter
Using licensed resources – not just copyright!

E-resources are generally licensed, rather than owned

<table>
<thead>
<tr>
<th>Contract law applies</th>
<th>Know the terms of your contracts, and be careful not to contract away your users’ rights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Example: text and data mining</td>
</tr>
</tbody>
</table>
Thank you!

Questions?

Contact:
Katie Zimmerman,
kbzimmer@mit.edu