

## The Google Library: 10 Questions



A Presentation  
by Siva Vaidhyanathan  
SV24@nyu.edu

---

---

---

---

---

---

---

---

## The Basics

- Google Book Search (formerly Google Print) is a service comprising two major projects.
  - A) The Publisher program -- a partnership with publishers through which they license their works to Google's index.
  - B) The Library program -- a partnership with libraries through which they offer millions of volumes for scanning into the index.



---

---

---

---

---

---

---

---

## Five Libraries



- Harvard
- Stanford
- Oxford
- New York Public Library
- University of Michigan
  - Only library that has publicly declared it will deliver post-1923 works to Google.

---

---

---

---

---

---

---

---

## The Copyright Issue

- Is the very scanning of the copyrighted works (post-1923) a violation of Sec. 106, the rights reserved for copyright holders?
- Does the fact that Google will only be presenting post-1923 works in “snippets,” much like KWIC searches via Lexis/Nexis, qualify the use as “fair” under Sec. 107?



---

---

---

---

---

---

---

---

## A Snippet

- Google promises that all works published since 1923 will only reach users as “snippets” so they will have no access to the entire text.
- This includes public domain works like US laws.



---

---

---

---

---

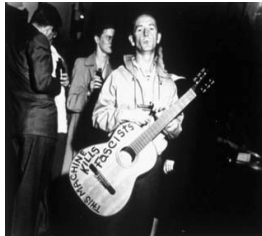
---

---

---

## Lessig's Fair Use Defense

- Precedent: Kelly v. Arriba Soft (2002).
- Fair use is supposed to work to enable public goods in the face of market failure.
- Most of the works of the 20th century have no identifiable copyright owner (orphan works problem).
- Therefore, it would be prohibitively costly if not impossible for Google to clear the copyrighted works.



Courtesy Woody Guthrie Archives

---

---

---

---

---

---

---

---

## UMG v. MP3.com (2000)

- MP3.com offered a “locker service.”
- Purchased thousands of CDs and made copies as MP3 files.
- Allowed registered users to put their own CDs in a drive, signaling MP3.com that the user owned the CD.
- MP3.com would allow access to the “locker” of music from anywhere.
- 2nd Circuit Court of Appeals in New York ruled the initial copying was infringing.



Duane Hoffmann / MSNBC

---

---

---

---

---

---

---

---

## New York Times v. Tasini (2001)

- Freelance contracts were for “first serial rights” only.
- Times and others sold electronic archive rights to Lexis/Nexis, ProQuest, etc. without securing electronic rights.
- Writers’ Union sued.
- Times: Impossible to negotiate rights now; public will suffer.

---

---

---

---

---

---

---

---

## East Coast vs. West Coast

- Both MP3.com and Tasini were “East Coast” cases in SDNY and Second Circuit.
- U.S. Supreme Court affirmed Tasini.
- Kelly is a “West Coast” case from Ninth Circuit.
- Kelly is a target.
- Google is imposing norms of the Web on to the real world.



---

---

---

---

---

---

---

---

## What Problems Might Google Book Search Solve?

- Limited access to book learnin'.
- Limited access to indexes and search functions.
- Book searches are guided by and limited by author, title, keywords, and subject headings.
- Obscurity of quality: Many good books are lost in the marketing shuffle.
- Inclusion of books to Web searches.

---

---

---

---

---

---

---

---

## Why Should We Care?

- Fair use is too rickety a structure to support so many essential public values and activities.
- Concern about rampant privatization of library functions.
- Concern about the frayed relationships between academic publishers and libraries.
- Frustration with bad Google search results.
- Frustration with my students' research techniques.
- Concern about user confidentiality for library use.
- Concern about the concentration of power in Google.

---

---

---

---

---

---

---

---

## What's at Stake?

- What matters more, the rights of established authors and firms or those of firms that push bold and radical plans the free-ride on the work of others?
- What matters more, the fact that Google would make an unauthorized copy of entire works for their own cache (an apparent violation of Sec. 106) or the interface that users will experience (a reliance on Sec. 107)?
- Have the processes of digitization and networking altered the assumptions of copyright, demanding a flexible and open vision of copyright?
- Will copyright cease to exist as a copy right and instead morph into a commercial distribution right, as scholars such as Jessica Litman have suggested?
- How should we address the market failure: By expanding fair use or by pushing for formalities?

---

---

---

---

---

---

---

---

## The Academy/Library in the Information Ecosystem



- Ignorance.
- Timidity.
- Confusion over role and mission: "The Content Provider Paradox."
- Michigan almost "gets it"

---

---

---

---

---

---

---

---

## Five questions for Google

- Google has been reticent about details.
- Google's goal: "to organize the world's information and make it universally accessible."
- Google's informal motto: "Don't be evil."

---

---

---

---

---

---

---

---

## Question One for Google

- What will be the guiding principles for inclusion, exclusion, and rank of the index?

---

---

---

---

---

---

---

---

**Question Two for Google**

- What will be scanned first? What order?

---

---

---

---

---

---

---

---

**Question Three for Google**

- What safeguards are you taking to ensure user confidentiality and privacy?

---

---

---

---

---

---

---

---

**Question Four for Google**

- What will be your metadata standards?

---

---

---

---

---

---

---

---

### Question Five for Google

- Will you omit certain titles from the index if a government requests it? Or will you merely present snippets to indicate the book's existence?

---

---

---

---

---

---

---

---

### Five Questions for the Libraries

- Libraries all signed contracts that included nondisclosure agreements.
- University of Michigan contract was subject to state open records laws.
- Unclear what libraries plan to do with the digital copies they receive from Google.

---

---

---

---

---

---

---

---

### Question One for the Libraries

- Did you insist on assurances that Google would protect user confidentiality and privacy?

---

---

---

---

---

---

---

---

### Question Two for Libraries

- Did you insist that Google's index include input from your librarians in terms of metadata standards?

---

---

---

---

---

---

---

---

### Question Three for Libraries

- Did Google restrict your use of the digital files in any way? (no obvious restrictions in Michigan contract).

---

---

---

---

---

---

---

---

### Question Four for Libraries

- Did you consider the harm to the potential markets for publishers who have been selling and leasing digital files?
- What is the copyright justification for receiving an electronic copy as payment for a transaction?

---

---

---

---

---

---

---

---

## Question Five for Libraries

- What's the hurry?

---

---

---

---

---

---

---

---

## The Myth of Neutrality

- Google is not neutral.
- Techno-optimism yields to technofundamentalism
- Google's search function is appropriate for the Web
  - Fast-growing, fast moving, ephemeral
  - Overtly intertextual
- Biases are built into its search systems
  - Opaque and proprietary

---

---

---

---

---

---

---

---

## Discussion Questions

- Are “snippets” useful as a resource for research?
- Play around with Google Book Search. Is it set up effectively? Does it serve users well?
- Do you think these libraries got a good deal out of Google?
- Do you think the public will get a good deal out of the Google-Library deal?

---

---

---

---

---

---

---

---