

# From Collections to Collaboration: Building for the Future

## A viability study of GreenGlass for analyzing the UC Berkeley Library collections

**Background:** OCLC’s GreenGlass is a tool for discovering duplicate monographs within OCLC’s own collection, or finding overlap with other institutions. A literature review of GreenGlass reporting suggests that it has most commonly been used for purposes of weeding.

### IN 2019, THE UC BERKELEY LIBRARY SUBSCRIBED TO GREENGLASS WITH THE FOLLOWING GOALS:

- Look beyond weeding and explore GreenGlass’ potential as a tool for collection development
- Test the viability of GreenGlass by developing case studies analyzing Berkeley’s collections
- Make recommendations for future use of the tool

### THE SCOPE OF THE BERKELEY HOLDINGS REPRESENTED IN GREENGLASS:

10.4 million print items | 6.2 million print titles

E-book data was problematic because of consortial cataloging and package issues, and therefore not included.

### GREENGLASS FEATURES:

- Customized peer groups for comparison purposes; for example, the University of California system, Big Ten Academic Alliance, UC Berkeley’s Standard Peer Comparison Group
- A variety of collection analysis parameters: subject, call numbers, location, record type, language, place of publication, possible duplicate, multi-edition title, multi-volume set, circulation, publication and acquisition date, HathiTrust availability, and a few others

### GREENGLASS’ STRENGTHS AND WEAKNESSES:

#### Strengths

- Useful for evaluating monographs and comparing holdings with other institutions
- Easy to navigate
- Productive for identifying unique collections
- Effective at identifying areas for collaborative collecting

#### Weaknesses

- Does not allow for comparison with specific institutions
- Data issues require post-processing and additional review
- Items without OCLC numbers cannot be integrated
- Not intended to be used for analyzing data on serials

### CASE STUDY 1 Would GreenGlass be effective at locating unique items in circulating collections that are frequently used?

This is a category of materials that is particularly vulnerable to damage and loss, the impact of which may be significant if these items are impossible or difficult to replace.

A query was designed to produce a list of **UNIQUE BOOKS FROM CIRCULATING COLLECTIONS** that are: **FREQUENTLY USED** **LIMITED TO CERTAIN LOCATIONS** **NOT AVAILABLE ONLINE**

**Step 1:** Limit the search results to books from circulating locations:

- Format = BOOKS
- Locations = Specific locations

**Step 2:** Limit the search results to books not digitized = NOT in HathiTrust.

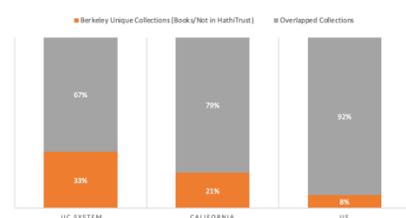
**Step 3:** For uniqueness, three geographic locations were used: unique within the UC system, unique within California institutions, and unique within U.S. institutions.

- UC system: fewer than 2 (any edition)
- California holdings: fewer than 2 (any edition)
- U.S. holdings: fewer than 2 (any edition)

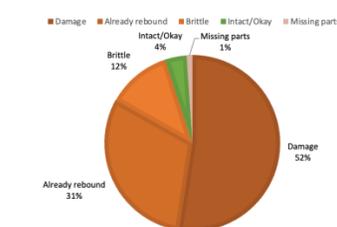
**Step 4:** For “high use,” three scenarios were used:

- Recorded uses = more than 2 or 5 or 10
- Last charge date = after 2015

**Chart 1.** Berkeley’s high-use and unique print collections within the UC system, California institutions, and U.S. institutions.



**Chart 2.** Physical condition assessment for a sample of unique titles.



#### Findings and recommendations:

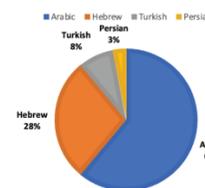
GreenGlass identified items that are unique and high use within three areas: the UC system, the state of California, and the United States. The proportion of those held by UC Berkeley dropped as the comparison group expanded. (See Chart 1.) Data-related complications emerged and necessitated manual review and culling of the unique title lists. This was caused by data errors and variations in cataloging; limitations of the tool; and the way the data was aggregated by the tool. A physical assessment of the items held at UC Berkeley’s main library (Doe Library) showed the majority to be damaged or vulnerable, confirming this category of collections to be high risk. (See Chart 2.) Further review is recommended to ascertain whether any have artifactual significance and warrant additional layers of protection. These being high-use items that are not available elsewhere in the country, researchers may benefit from digitization. Items that fit into this category of materials may also be used to divine trends that inform collection development, exhibit ideas, or “stories” for the Library to tell about how the collections are used.

### CASE STUDY 2 Can GreenGlass be a useful tool for collaborative collection development?

Does GreenGlass reveal areas of overlap that UC Berkeley has with other institutions? Can areas of overlap be used to inform collecting strategies? Can potential partners be identified for building shared collections? Knowing how much collections overlap with those at other institutions is critical for identifying potential partners to build shared collections.

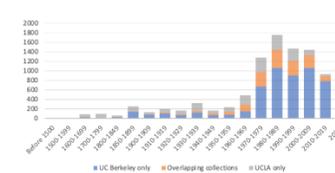
**THIS CASE STUDY** examines three area studies collections at UC Berkeley — Slavic, Middle Eastern studies, and Japanese Buddhism — and compares them to those held at UCLA. The GreenGlass search results revealed areas of overlap between UC Berkeley’s collections and collections of other institutions, but did not identify the specific UC campuses or institutions with which there is overlap. This limitation was overcome by using the WorldShare Collection Evaluation service (developed by OCLC) in conjunction with GreenGlass.

**Chart 3.** Language distribution of the noncirculating overlapping Middle Eastern studies monographs at UC Berkeley and UCLA.



A deeper analysis of the noncirculating Middle Eastern studies collections at UC Berkeley and UCLA reveals the degrees to which materials in various languages overlap. For instance, the pie chart shows 61 percent of the overlapping noncirculating Middle Eastern studies monographs are in Arabic. Focusing on the areas where the collections overlap highlights the opportunity for collaborative collecting and minimizing duplications of potentially low-use materials.

**Chart 4.** Overlapping Japanese holdings in Buddhism between UC Berkeley and UCLA by publication year.



Further deepening the analysis in a given subject, UC Berkeley’s Japanese holdings in Buddhism were compared against UCLA’s holdings. To identify overlap between the institutions, this analysis also used collection data from the WorldShare Collection Evaluation service. The analysis not only gives the breakdown of the data by publication year, but also shows specific overlap in both libraries’ Japanese holdings in Buddhism.

#### Findings and recommendations:

GreenGlass is a useful tool in helping to identify areas of collection strength when comparing against peer libraries. This makes it very useful beyond collection weeding. A collection’s strength depends on a number of factors, such as where and in what languages its materials were published, and nuanced search features in GreenGlass enable a more precise evaluation of area studies collections. The area studies collections used for this case study were made further complicated by their interdisciplinary nature and the variety of formats and languages. A simpler use case focusing on fewer collection variables would be easy to prepare by comparison. While GreenGlass on its own cannot be used to discover a potential partner, combining it with other tools such as OCLC’s WorldShare Collection Evaluation service is an especially powerful strategy for discovering potential partners for collaborative collection development.