

Google Books' Coverage of Hawai'i and Pacific Books

Andrew Weiss | California State University, Northridge & Ryan James | University of Hawai'i at Mānoa

The *Google Books Digitization Project* is one of the world's largest online massive digital libraries. But how does it fare in terms of diversity of subject and traditionally under-represented groups? This study examines *Google Books'* coverage with respect to a collection of Hawaiian and Pacific books housed at the University of Hawai'i at Mānoa in comparison to works found in its library's general stacks collection.

1. METHODOLOGY

- **STEP 1** 3 random samples of 500 book records each were taken from 3 UH-Mānoa collections: the Hawaiian, the Pacific, and the general stacks collections. **1,500 records in total.**
- **STEP 2** Selected records were then **searched for in Google Books.**
- **STEP 3** Matching item records found in Google Books were **evaluated for their level of access.** The categories used were 'Metadata Only,' 'Snippet,*' 'Preview,*' and 'Full View'.
- **STEP 4** Selected item records were **recorded at their highest level of access.** Selected item records **not found in Google Books** were recorded in the **"No record"** category.

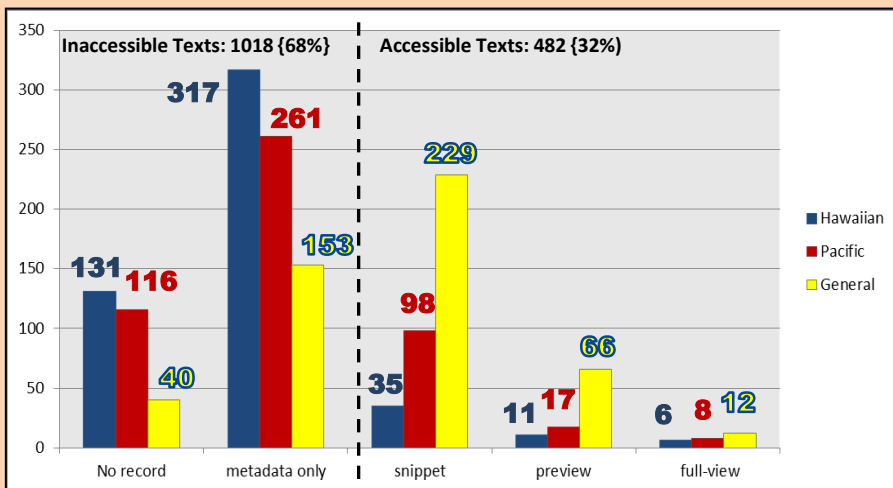
*NOTE: Snippet = a few sentences are accessible;
Preview = a few pages are accessible

2. RESULTS

- **Hawaiian collection books (N=500):**
 - 131 had no record in Google Books,
 - 317 had a metadata record only,
 - 35 had a snippet view available,
 - 11 were available for preview,
 - 6 could be fully viewed.
- **Pacific collection books (N=500):**
 - 116 no record,
 - 261 metadata record only,
 - 98 snippet view,
 - 17 preview,
 - 8 could be fully viewed.
- **General stacks collection (N=500):**
 - 40 had no record,
 - 153 had a metadata record only,
 - 229 had a snippet view,
 - 66 were available for preview, and
 - 12 could be fully viewed.

3. DISCUSSION

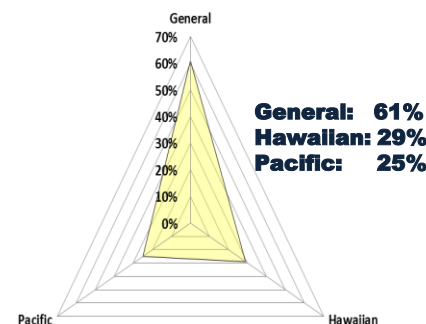
- **68%** of sampled books have no record in Google Books or have only metadata records and are not accessible as texts, suggesting that these books have not been digitized;
- **Accessible or partially accessible texts represent only 32%** of the collection, again suggesting lack of digitization;
- **Of those, 66%** of books accessible with at least a snippet view come from the General Collection, likely covered by Google's main partners on mainland USA library collections;
- **Only 175 of 1000 texts (17.5%)** from the Hawaiian and Pacific collections are accessible with at least a 'snippet' view;
- **Of books in the public domain, 61%** in the General Collection are accessible; significantly less for Hawaii and Pacific Collections..



Items sampled and their corresponding levels of availability in Google Books.

Subset: Public Domain Books

% of Public Domain Books Available in Full View



4. CONCLUSIONS

- **Acute differences in accessibility in Google Books between a general collection & Hawaii/Pacific books;**
- **Most books in H/P collections are likely not digitized; a result of Google's partnering primarily with mainland U.S. universities;**
- **Further studies of Public Domain books and their accessibility in Google Books are warranted;**
- **Further studies of accessibility for other traditionally under-represented groups are warranted;**
- **Comparative studies of accessibility in Google Books, Internet Archive, and the Hathi Trust are warranted;**