## U-M University Library – Google Digitization Project: A Brief Overview Dec. 13, 2004

The University of Michigan and Google, Inc. have entered into a partnership to digitize the entire print collection of the University Library. The digitized collection will be searchable by Google, and the University Library will receive and own a copy of all images to integrate into new and existing U-M Library user services.

The U-M Library embarks on this ground-breaking partnership for a number of very compelling reasons.

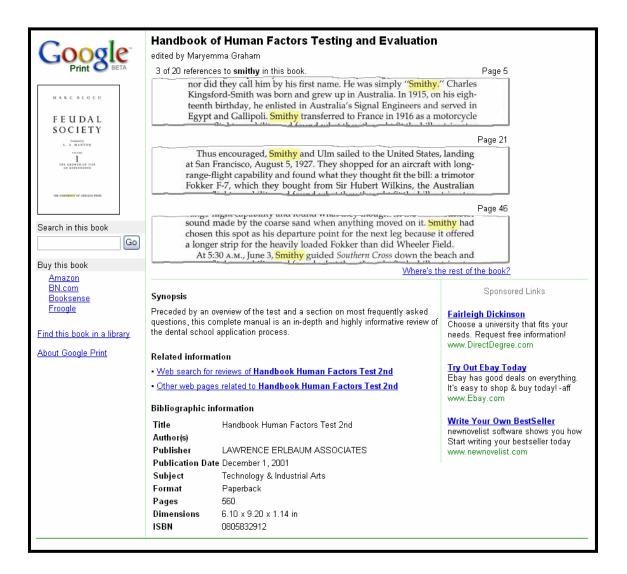
- Most significantly, this work will create new ways for users to search and access library content, opening up our collections to our own users and to users throughout the world.
- Although we have engaged in large-scale (preservation-based) conversion of parts of the Library's collection for several years, we know that only through partnerships of this sort can something of this scale be achieved.
- We believe that, beyond providing basic access to library collections, this activity is critically transformative, enabling the University Library to build on and reconceive vital library services for the new millennium.

In 2002, Google began discussions with the University Library to digitize their collection. Google has similar partnerships with Harvard and other universities, though U-M is the premiere testing site for Google's non-destructive scanning technology and digitization workflow, and is the first site to implement such technology.

Work has begun to convert materials at the University of Michigan. Nevertheless, because of scaling and engineering issues, it will be many months before materials begin to appear in Google's services, and approximately six years until the collections at the U-M University Library have been converted in their entirety.

Google will add the resulting millions of volumes of digital content to its Google Print service, which currently includes print volumes converted in cooperation with publishers. Once the library content is integrated into Google Print, users of Google's services will be able to find books and journals as part of their general searches. There will be **no charge** for access to the library content. It will be possible for a user to search the full-text of all the volumes digitized and to **view every page of an out-of-copyright book**. A Google search returning results from a book **in-copyright** will return (see Figure 1):

- three "snippets" of text from throughout the volume, a "snippet" consists of approximately 3 lines of content;
- a count of the number of times the search term appears in the volume,
- the bibliographic data associated with the book; and
- information on where to buy the book **OR** find it at a local library.



The terms of the agreement ensure the following for the University of Michigan:

- The University Library is provided a high quality digital copy of the materials. The quality of the capture and digital files is such that the digital files are consistent with library preservation community standards for readability and best practices. The processes Google uses to digitally capture the books and journals are non-destructive and do not disbind Library materials. The methods Google uses for digitization resemble ordinary (though careful) use of library materials; once they are scanned, books and journals are returned to Library shelves.
- The University Library is not limited in the use of our digital copies to build services that meet the needs of users in the University and larger academic community and is wholly committed to ensuring the digital version's long-term preservation. The Library's use of these files will be within legal and mission-oriented parameters.

The terms of the agreement ensure the following for Google:

- Google may use the files in the delivery of information as part of the array of Google services.
- Google must work within the constraints of copyright laws, and to that end will only show a small amount of information from in-copyright materials (e.g., up to three 3-line "snippets" of content from a given in-copyright book).

Larry Page, the President and Co-Founder of Google, and an alumnus of the University, is credited with first approaching the University with his idea. The leadership of the Executive Council was pivotal in bringing this project forward to the University Administration. Behind the scenes work included confidential discussions and proposals with Google staff, U-M legal counsel, and strategic staff within the Library and University. All parties worked together to carefully sift through serious issues and concerns involving copyright, the rights of the each partner, distribution of responsibilities, and digitization standards and best practices. The outcome is a remarkable project which will enhance the University Library and the services it provides its users well into the 21<sup>st</sup> century. It will change the Library landscape in ways we have not fully imagined.

For further information:

http://www.lib.umich.edu/staff/google

http://print.google.com/library