Judy McConarty

ILS655 Assignment-Digital Library Project Review Paper 01

February 6, 2011

Project Gutenberg

1. Background

Project Gutenberg (PG)'s mission statement is simply "To encourage the creation and distribution of books." It is the oldest digital library, founded in 1971 by Michael Hart when he was a student at the University of Illinois. At the time, he was provided access to a powerful mainframe computer and given a considerable amount of computer time in the university's Materials Research lab. He decided to search for public domain texts (those no longer protected by copyright), digitize them and build a collection available to everyone free of charge. The first text he digitized was the Declaration of Independence. The organization relies on donations for financial support and the collaborative efforts of volunteers to digitize text, proofread and perform other tasks. As a result of extensive volunteer collaboration and improvements in technology, the library's rate of growth has increased over time. As of the end of 2010, the collection included roughly 39,000 eBooks. PG Literary Archive Foundation is the not-for-profit corporation that receives donations, fundraises and otherwise manages the organization ("About," 2008). The URL is http://www.gutenberg.org. The Website is hosted by iBiblio.

2. Organization of Resources

PG has digitized roughly 39,000 eBooks, most of which are in the public domain. Approximately 29,000 of these are in English. Remaining texts are in French, Italian, Chinese, German, Spanish, Portuguese, Greek and other languages. Texts are not actually held on the Project Gutenberg Website but on FTP (File Transfer Protocol) sites around the world. The Website actually just holds an index ("Readers'," 2010). PG's three main eBook categories are light

literature (e.g., *Peter Pan*), heavy literature (e.g., the *Bible*) and references (e.g., *Roget's Thesaurus*).

Beyond eBooks, sub-projects include the Audio Book Project, the CD and DVD Project and the Sheet Music Project. PG is affiliated with several other projects of independent organizations that share the same philosophy (e.g., PG of Europe, PG of Australia). These organizations are entirely separate but allowed to use the PG trademark ("General," 2008).

Metadata methods include the use of different schemes. A PG scheme is used for data relevant to PG such as eBook number. A Dublincore scheme is used for data used by search engines and cataloging software. Additionally, a MARCREL scheme is used to record contributors ("Metadata," 2010).

3. Service Features

The online catalog is user-friendly and can be accessed without cost or registration. It allows for a variety of search options, as shown below.

An advanced search feature is also available, which provide additional search options such as file type, item category, Etext number, etc. The "Bookshelves" feature facilitates browsing. Categories such as best books ever, banned books, children's literature, philosophy and others can be selected to browse desired genres/subjects within the collection. Authors, titles, subjects and bookshelves can be sorted alphabetically, by popularity or by release date. Other search engines provided are Anacleto, Yahoo and Google. Anacleto performs full-text and metadata (author/title/keyword) searching. Yahoo performs metadata searching and Google performs full-text searching.

4. Technologies

The catalog and eBooks can be viewed with a Web browser and texts can be downloaded in various formats, including ePub, Kindle, HTML and simple text. ASCII (American Standard Code for Information Interchange) is the preferred format because almost any hardware and software a user might run can read and search these files. This is consistent with the organization's mission to make eBooks available in the simplest, most accessible form possible. Texts can be read on PCs, iPads, Kindles, Sony Readers, iPhones, Androids and other devices.

5. Comments

I'm fascinated by how PG has evolved over time. From the digitization of the Declaration of Independence in 1971, it has grown to a collection of 39,000 eBooks, not to mention initiatives in other formats. Based on my experience, PG is living up to its mission "to encourage the creation and distribution of books." I found the library well organized and the catalog user-friendly. It is easy to search or browse the collection and sort results. I also tried downloading an eBook and was pleased to find that it was really simple. Not surprisingly, *School Library*

Journal included PG in a list of recommended online reading resources (Adam, 2008).

Concerns about the validity and reliability of user-generated content are addressed in an article in *Searcher* by Nicholas Tomaiuolo, Instruction Librarian at Central Connecticut State University. While acknowledging that not everyone shares the same opinion, he says it is obvious that user-generated content is not going away, as evidenced by the popularity of Wikipedia, Facebook and other sites. As such, the focus should turn to making the content better (Tomaiuolo, 2009). This is an ongoing effort for PG, whose volunteers continually proofread text and correct errors when they are found.   In a recent interview, Michael Hart described his view of PG's use of the Internet this way:

"I see Project Gutenberg's use of The Internet as the primitive, first steps sort of *Star* Trek Communicator, Transporter, and in the most important way, as a very primitive "Replicator." You put a book in over here, and anywhere on The Internet every person can have a copy, free of charge. Today "The Personal Computer" Becomes "The Personal Library."The average computer today goes for under $500. For another $500 you can add your first few terabytes, and hold enough books (in plain text.zip files] to make it into current lists of the top 100 largest libraries INTHE WORLD!!! " (Interview, 2009)

References

Adam, A., & Mowers, H. (December 2008). The Inside Story: Don't stop at the cover. *School Library Journal. 54*(12), 50-2. Retrieved on February 4, 2011 from First Search.

About, 2008. Retrieved on February 3, 2011 from:

http://www.gutenberg.org/wiki/Gutenberg:About

General FAQ, 2008. Retrieved on February 2, 2011 from:

http://www.gutenberg.org/wiki/Gutenberg:General_FAQ

Interview With Michael Hart. (January 2009). *Searcher. 17*(1), 32-3. Retrieved on February 5, 2011 from First Search.

Metadata Directive, 2010. Retrieved on February 5, 2011 from:

http://www.gutenberg.org/wiki/Gutenberg:The_PG_boilerplate_for_RST

Readers' FAQ, 2010. Retrieved on February 4, 2010 from:

http://www.gutenberg.org/wiki/Gutenberg:Readers%27_FAQ

Tomaiuolo, N. (2009). U-content: Project Gutenberg, me, and you. *Searcher*, 17(1), 26-34. Retrieved on February 4, 2011 from EBSCO*host*.