

Google Book Search Libraries and Their Digital Copies

by *Jill E. Grogg*, Electronic Resources Librarian, The University of Alabama Libraries
and *Beth Ashmore*, Cataloging Librarian, Samford University

Few things in the past decade, other than the PATRIOT Act, have brought libraries and subsequent controversy into the mainstream media as much as the Google Book Search Library Project.

For some — both inside and outside the profession — the mass digitization of library-owned books by Google sounded like yet another death knell for physical libraries and their custodial librarians alike. For others, it appeared to launch the mother of all copyright cases. However, in nearly every instance of media hype, the focus sat squarely on what Google planned to do with all those digitized books. While Google's intentions are always a good topic for conversation and as everyone waits for the courts to decide on important copyright issues, one can't help but wonder: How will the librarians at participating Google Book Search libraries use their copies of the digitized books, commonly referred to as the library digital copy, the copy that Google gave to them in return for their participation in the Book Search project?

Google Book Search participating libraries now include the University of California (UC), University Complutense of Madrid, Harvard University (Harvard), University of Michigan (UM), The New York Public Library (NYPL), Oxford University (Oxford), Stanford University (Stanford), University of Wisconsin-Madison (UW-Madison) and University of Virginia (UVA). In January 2007 alone, Google added two more library partners with the National Library of Catalonia merged with four affiliate Catalonian libraries and the University of Texas at Austin library. In February, Google added Princeton University's libraries and, in March, Bayerische Staatsbibliothek (Bavarian State Library). These libraries represent some of the largest and most impressive collections in existence anywhere in the world, and, in most cases, the deals that have been struck with Google are just as unique as the collections being housed. Some of the library administrators have chosen to focus on specific collections within their libraries, while others simply look for materials in good condition and in the public domain. Speaking of public domain, many library administrators have chosen to restrict themselves to these materials. Others either leave the door open to any materials within their collection, public domain or not, or purposefully choose to include in-copyright materials.

Once the contracts were signed, the scanning began at full speed. However, it is critical to note that while Google scanning has begun, partnering with Google is not the first experience these librarians have had with digital preservation and full-text searching of their digital content.

Pre-Google Library Digitization

The Google Book Search Library Project is extremely attractive because Google has the knowledge and the resources to shift a library's existing digitization program into high gear at the best possible budget

allocation — nearly free. However, considering the immense and unique collections of the Google Book Search Library Project partners, it is no surprise that many of these libraries had thriving digitization projects underway long before Google came knocking.

University of Michigan, arguably the leader among the Google library partners, has been working on a variety of digitization initiatives since the late 1980s. Anne Karle-Zenith, special projects librarian, University Library, UM, cited the current statistics on their digitization progress prior to Google: “141 text collections with 25 million page images online, plus 3 million pages of encoded text and 89 image collections containing approximately 200,000 images.” UM has also partnered with Cornell University to create the Making of America project, funded by the Mellon Foundation. Making of America has provided researchers with access to hundreds of volumes of American primary sources from 1850 to 1876. The established reputation of UM's library as a leader in digitization merely represents an outgrowth of its mission. Karle-Zenith explained: “Preservation and stewardship of our digital assets has always been one of our top priorities.”

At New York Public Library, The Digital Gallery contains more than 520,000 images from the four research libraries: the Humanities and Social Sciences Library; The New York Public Library for the Performing Arts; the Schomburg Center for Research in Black Culture; and the Science, Industry and Business Library. David Ferriero, Andrew W. Mellon director and chief executive of the NYPL Research Libraries, described The Digital Gallery as “... a kind of snapshot of collections from within the four research libraries.” Another NYPL digitization project, In Motion: The African American Migration Experience, comes from the Schomburg Center for Research in Black Culture with support from the Congressional Black Caucus and the Institute for Museum and Library Services. NYPL has other digitization efforts, all available under the moniker NYPL Digital, at <http://www.nypl.org/digital/index.htm>.

UW-Madison, a recent addition to the list of Google Library partners, is no stranger to digital preservation. Edward Van Gemert, interim director at the UW-Madison Libraries, explained: “We've collected, digitized, organized, and made available now close to 2 million pages of content with a full range of subjects and all of that material is available on our library Web site.” UW-Madison's current digitization projects, including its plans with Google, have a well-defined focus born directly out of the strength of its American history collections and its ongoing partnership with the Wisconsin Historical Society (WHS). By working with state and federal documents and other public domain materials that cover areas including “statehood, regional history, patents and discoveries,” UW-Madison seeks to create a historical record of the formation of the U.S., the upper Midwest, and the territory of Wisconsin. By playing to the strengths of its physical collections, UW-Madison seeks to establish a digital repository of primary resources in American history. The partnership among UW-Madison, WHS, and Google is designed to build upon this existing mission.

UM, NYPL, and UW-Madison are by no means alone in their pre-Google digitization endeavors. According to the University of California 2006 annual report [http://www.universityofcalifornia.edu/annualreport/2006/pdf/fullreport_06.pdf], “Calisphere, an online service of the UC Libraries and the California Digital Library, provides access to over 170,000 digital images and 50,000 pages of documents about California.” The UC Libraries and the CDL have other digitization projects as well.

Dale Flecker, associate director of the Harvard University Library for Planning and Systems, noted, "Harvard has a strong and long-standing program in preserving digital information. We will use our digital preservation infrastructure to preserve the data created in our Google project." Michael Popham, head of the Oxford Digital Library, Oxford University Library Services, echoed Flecker's comments about long-standing programs: "We [Oxford] have undertaken a variety of digital preservation initiatives over many years across the university. For example, this year the Oxford Text Archive celebrated its 30th anniversary collecting, preserving, and freely disseminating electronic texts and corpora."

Finally, the University of Virginia is a long-time player in digital preservation. Karin Wittenborg, university librarian, UVA, noted, "We started, I think, the first electronic text center in the humanities back in 1992, and so we've been digitizing for a long, long time, and we know exactly how expensive it is." This digitization effort was originally called Etext but, according to Wittenborg, UVA has merged many texts and images into something it calls Scholars' Lab [<http://www.lib.virginia.edu/scholarslab>]. In the original Etext Center [<http://etext.lib.virginia.edu/collections>], items scanned were all public domain materials.

At the end of the day, the participating libraries' existing projects represent an effort to achieve two key goals: to preserve materials for generations to come and to provide increased access and functionality for the generation at hand. With Google the favorite discovery tool among the current generation, it is easy to see how Google has become an important partner for libraries to further their digitization goals.

The Google Library Party

Librarians participating in Google Book Search scanning can easily add links on their libraries' Web pages to the Google Book Search Web site [<http://books.google.com>] and call it a day. However, with the number of aforementioned in-house digitization projects, most librarians find great potential for using the copies they receive from Google in conjunction with existing library resources and newly created partnerships. And the future may see even nonmembers of the Google Library Party doing the same thing. As Megan Lamb, of Google Corporate Communications, explains "We've [Google] already done significant engineering work in ensuring that our URLs are persistent and any organization can link to them."

For example, Popham of Oxford explains the necessity for new partnerships to cope with the enormity of the Google project: "The sheer scale of our endeavor with Google vastly overshadows any previous [digitization] activity and will require additional preservation infrastructure, which we are developing in partnership with Sun Microsystems, as part of the establishment of a Sun Center of Excellence here in Oxford." While all of the participating Google Book Search libraries have digitization projects in existence, few, if any, approach the scale of the Google project. Jennifer Colvin, strategic communications manager, UC Office of the President, stated, "The university libraries have been doing other digitization projects for years, but nothing nearly on the scale of what we are doing with the OCA [Open Content Alliance] and with Google, so that's one of the reasons we are so excited about our partnerships with those organizations."

Such an increase in scale means that some library administrators are still weighing options about how to use their library digital copies. In August 2006, Barbara Quint reported in an Information Today, Inc.

NewsBreak ("Google Book Search Adds Big, Brave Partner: The University of California") that "plans as to what UC intends to do with its digital copies are still in the works. However, public domain material will have free and unfettered full-text access throughout the system, including links to the online Melvyl Catalog. Books still in copyright will only be accessible in keeping with copyright law"

[<http://newsbreaks.infotoday.com/nbReader.asp?ArticleId=17375>]. According to Colvin, UC has organized a "system-wide group with representatives from across the UC system to try to figure out what the next step is going to be and how we can possibly integrate those digital books in with our collection." According to Flecker, Harvard is not using the data at this point: "Future uses are under discussion, but no concrete plans are in place." While no concrete plans may be in place, Harvard is enthusiastic about the possibilities the mass digitization offers. Flecker noted, "We are excited about the possibility of making the collection of scanned books available in the future for text mining, which we believe will open up powerful new ways of doing research."

NYPL's Ferriero envisions a future in which patrons conduct searches in its Digital Gallery and receive not only images from across the research libraries, but also the text that provides necessary context and avenues for further research. While NYPL is still making plans for how it will use its digital copies (in December, it just finished the Google pilot of 10,000 volumes and made arrangements to continue the relationship), it is carefully watching how other Google partners are putting their digital copies to work for them. In fact, the Google Library partners meet twice a year to take advantage of the lessons learned from each of their very individual ventures.

Like Harvard and the NYPL, Oxford continues to explore how best to use its digital copies. When asked to describe how his library currently uses and/or plans to use digital copies received from Google, Popham replied, "At the moment, we are simply planning to archive and preserve our copy of the data generated by our joint project with Google." Popham went on to say that Oxford will link from its catalog record to the images hosted at Google [<http://books.google.com>]. Finally, Popham explained, "The scale and scope of this project is such that we are only just beginning to consider some of the possibilities that this work may enable."

As one of the most recent of Google library partners, signing on in November 2006, UVA is understandably still considering how to use its digital copies. When asked if UVA was concentrating on specific subject areas, Wittenborg said, "We gave Google a lot of data on what we think our special strengths are, but they said essentially all 5.1 million volumes are under consideration still by them." Wittenborg went on to say, "They [Google and library partners] have a summit meeting of the [library] partners in January [2007] at Google, and so I think by then, if not before, we'll know. We are very, very strong in American literature and American history, but also in Buddhism and other things, so it's pretty much they get to choose." When asked if UVA was choosing to do both public domain and in-copyright materials, Wittenborg replied, "Absolutely. It was important for us to want to do the whole thing." Wittenborg emphasized the opportunity this presents for UVA: "... it's an open playing field. We know we'll be experimenting with software tools and delivery services, and our primary goal is to support research, teaching, and learning here at UVA. But I think that once we suddenly get content, we will find out there are all kinds of things we can do. I think there will be parts of the content that we will mark up for added value, but we just don't know yet."

UW-Madison, another new member of this exclusive club, has particular plans for organizing and providing access to the library's digital copies of Google-scanned material. Van Gemert stated: "Our intention is to have material searchable through our OPAC and our intention is to collaborate with other CIC [Committee on Institutional Cooperation] institutions on a shared digital repository." The CIC is a consortium of 12 major research universities, including those from the Big Ten athletic conference, along with the University of Chicago and the University of Illinois at Chicago. As mentioned earlier, UW-Madison also intends to leverage its investment in current digitization projects and, through its partnership with the Wisconsin Historical Society and Google, to create a larger, more comprehensive digital collection. Van Gemert explained, "Our primary focus is to have digitized materials in the public domain — state and federal government documents, other historical documents." While UW-Madison has a clear focus, it has yet to begin scanning; the operational side of its project with Google will begin in March 2007.

In addition to UW-Madison, the CIC group also includes UM, which, to date, has one of the most developed systems for providing access to its Google scanned materials — MBooks. MBooks allows patrons to discover books through full-text searching in its online catalog, Mirlyn. Once a title is identified, a patron can click on a link, which takes them to a "page turner" interface that allows them to navigate the book, print individual pages, and enlarge and rotate the page image as well as to search within that individual title. The library digital copies delivered to UM are individual pages — mostly 600 dpi TIFF images using ITU G4 compression, although pages with significant illustrations usually appear as 300 dpi JPEG 2000 images. Google also provides an OCR text file to match each page image. The type of image files provided to a library depend on the library's preferences [<http://www.lib.umich.edu/staff/google/public/faq.pdf>]. While one cannot print out entire PDF files of a single title from the MBooks interface, the MBooks record does provide links back to the Google Book Search copy.

Because UM is one of the libraries allowing Google to scan in-copyright titles, MBooks provides searching within these titles, as well as information on the number of occurrences and location of terms within an individual title, to assist the user in evaluating the relevance of the item to their research. MBooks currently only houses the titles scanned through Google, but there are plans to add materials from previous in-house digitization efforts to the database as well.

Public Reaction

The media has definitely taken a shine to the Google Book Search Library Project. However, Google is not the first to try to take the world of public domain online. Initiatives such as the Internet Archive, Project Gutenberg, and American Memory have all tread much of the same ground as Google. So, why all the fuss? Some of the attention can be traced back to Google's high recognition factor, which inevitably makes its new endeavors newsworthy. And then Google is clearly navigating some uncharted waters for fair use. The speed and scope of the effort, covering such a large amount of materials in a relatively short amount of time, also draws attention.

Karle-Zenith summed it up: "This is a very ambitious project that will provide scholars and the general public with an unprecedented ability to search for and locate books from the university's vast collections. This initiative has the potential to revolutionize the way the world's knowledge is transmitted and to democratize access to information. However, throughout history, breakthroughs in technology have

always created challenges." Wittenborg also commented about the media fascination with Google: "I think the media is captivated because Google is changing the game. The rest of us were moving very slowly with public domain ... and suddenly everything is different. The potential for discovery at this level of magnitude of millions of titles is going to be incredible."

While the media may be taking notice, users still seem to be a bit confused over what exactly this project means for their own research. "I think there is a fair amount of concern over what this all means for the future," Ferriero explained, adding that he thinks there is a great deal of excitement about the ability to search full text. "The fact that the New York Public Library has decided to include public domain material only is different from some of the other partners, and I think that's not clear in some people's minds, so they expect to see content that hasn't been digitized."

If the differences between the arrangements that Google has made with each of its library partners have led to confusion on the part of users, it has not prevented new library partners from seeing the advantages to signing up. "Campus administration was clearly supportive and clearly excited about the project," Van Gemert said of UW-Madison's recent negotiations with Google. "We see it, of course, as a way to get a huge amount of material digitized for a relatively low cost."

The library administrators were nearly unanimous in responding that their respective constituencies were supportive of the project. Popham of Oxford noted, "On the whole, the reaction had been extremely positive." However, similar to Ferriero's comments about general confusion, Popham said, "Most of the reservations that have been expressed to us have been based on a misunderstanding of the nature of the project, or the way the digitized materials will be made available. For example, the most common misconceptions are that access to the digital copies of materials from our collections will be restricted to users at Oxford, or that people will have to pay to access the digital materials; in fact, neither is the case." Therefore, while the media has showered much attention on certain aspects of the Google Book Search Project, it will still be incumbent on librarians at participating libraries to educate their users about the realities of participation.

Copyright and Other Concerns

In any discussion of Google Book Search, the 800-pound gorilla in the room is copyright, or, more accurately, litigation over copyright. In the fall of 2005, both the Association of American Publishers (AAP) and the Authors' Guild filed suits against Google over Google's library scanning project. Jim Milliot summed up the major issues behind the AAP's suit in an Oct. 24, 2006, column in *Publisher's Weekly*: "The lawsuit reflects a deep division between publishers and Google over the meaning of fair use. Google compares scanning books to its copying of materials online, a comparison the publishers contend is faulty." The fact that there is so much ambiguity over what constitutes fair use will come as no surprise to information professionals, but the serious way in which a critical mass of the publishing community has banded together on this issue is noteworthy. "Google has positioned its library project as part of its mission to make all the world's information available," Milliot explains, "but publishers see scanning of copyrighted materials without permission as the first step in the loss of control over their content."

It appears that for copyright holders, the stakes could not be higher, making the likelihood of publishers backing down fairly low. No doubt, the library administrators involved in the scanning of in-copyright

material have taken this copyright smackdown into consideration before signing on with Google. As a matter of fact, one library declined comment for this article partly due to the pending litigation. When asked to be interviewed, Michael A. Keller, university librarian; director of academic information resources; publisher of HighWirePress; and publisher of Stanford University Press, Stanford University, noted that Stanford preferred to comment about its intentions for its digital copies “only after we have actually accomplished some of the programs and functions that we anticipate implementing.” Keller further explained, “It would be a mistake for anyone at Stanford involved in the project to respond to some of your questions before the suits against Google in the Southern District of the Second Circuit of the U.S. courts have been heard and decisions made and publicized.” Should readers be interested in Stanford’s participation, Keller pointed to a number of Web sites, such as http://library.stanford.edu/about_sulair/news_and_events/stanford_google_project.html.

With both Google and publishers seemingly equally convinced of their legal footing in the ensuing battle over fair use, most library administrators appear to be taking a safe approach to scanning. For many years, librarians have been trying to discern the best way to comply with copyright while providing as much access as possible. It is not easy for librarians to let go of years of cautious interpretation of fair use, but some are making an effort. UM’s Karle-Zenith stated, “We believe Google is making a lawful copy.” However, she also noted, “The library is assuming little to no risk because, per our contract, Google indemnifies us against any third-party claim that the project violates third party’s copyrights or other legal rights.”

Even with UM’s relatively safe approach to in-copyright scanning, many library partners are still going the public domain route — for more than simple legal reasons. “We wouldn’t get involved until there’s a decision on the two lawsuits,” NYPL’s Ferriero explained. “The rationale is that we’re here in New York City in the middle of the corporate publishing empire, and many of those publishers’ names are carved in gold on our walls. So, we are nervous about getting into that arena.” In addition to trying to keep up amicable relations with publishers, some library administrators see public domain as a big enough task already. “There’s a huge amount of material in the public domain ... there’s a whole category of material that falls into that orphan class of material that is indistinguishable ... so that is an area where we will probably move into next,” UW-Madison’s Van Gemert explained. With a world of pre-1923 titles waiting for attention and legal decisions yet to be decided, the hassle of in-copyright scanning would appear to be too much for most participating libraries to attempt — at least for now.

Partnering with Google on any project carries with it some concerns of its own. Brewster Kahle, founder of the Open Content Alliance (OCA), has expressed concerns that Google is building “the private library” of a single corporation rather than a public resource, and questions whether or not this is the right kind of project for libraries to become involved in (*Library Journal*, Oct. 1, 2006). Many library partners see this not so much as an opportunity to help Google build its own library, but instead a situation where both parties reach their goals through collaboration. “The University [of Michigan] has a long history of collaborating with private corporations when we can find areas of common interest and can work together to produce fruitful outcomes on both sides,” explained Karle-Zenith. Other library partners seem to take a different tack, participating in many digitization efforts at once. For example, UC is a member of the OCA, Google Book Search, and Microsoft Live Search Books, as well as other digitization initiatives. When asked about UC’s decision to participate in multiple initiatives, Colvin replied, “I think

this goes back to our mission as a public university and our goal to make as much information available as possible. We are happy to work with anyone who will help us achieve those goals." Like UC, UVA is also a member of the OCA.

Still other librarians feel it is necessary to take steps to ensure that Google's program is in line with their library's mission. As Van Gemert explained, "Anytime you go into a large project like this that represents significant change ... there's going to be concern. Some folks definitely want to talk about what safeguards we've put in place to obey copyright, to maintain a research library in terms of collection, whether they be print or electronic, and proper preservation activities..."

The concerns don't end there. With so many partners sticking to their public domain collections, there is some question of the value of these materials to the average user. In his Feb. 5, 2005, *Library Journal* column, Roy Tennant questions whether or not we are doing users any favors by adding a plethora of pre-1923 information to the already authority-shaky Internet: "Unfortunately, I can think of few situations where having access to only pre-1923 literature is a good thing. The typical user who finds a pre-1923 source available for free via Google is unlikely to sashay down to the local library for something more recent. That's just life." While many information professionals would echo Tennant's sentiments, Karle-Zenith sees this as no reason to dump the project: "How could access to only pre-1923 literature be better than access to no literature? We have seen the value of pre-1923 content time and again, with, for example, the Making of America [MoA] project.... The response from scholars has been enthusiastic from the time the MoA materials first went online.... To this day we continue to hear from users about new discoveries and new knowledge generated by their research on Making of America." Like the enthusiastic scholar reaction to MoA, UVA's Wittenborg commented on the "extraordinary number of downloads" from UVA's Etext Center, which contains, in addition to other items, a number of digital copies of 18th- and 19th-century books. "Suddenly," Wittenborg said, "these books found an audience."

Furthermore, some librarians and university administrators have expressed that partnering with Google is simply "the right thing to do." Wyatt R. Hume, UC executive vice president and provost, expressed precisely this sentiment in an Aug. 9, 2006, press release [<http://www.universityofcalifornia.edu/news/2006/aug09.html>]. In the same press release, Brian E. C. Schottlander, university librarian at UC San Diego, rationalized that participation with Google is a solution to urgent problems of preservation, giving Hurricane Katrina as a prime example of the type of natural disaster that can wreak havoc.

Whatever the reason for participation, whatever the rationale and accompanying concerns, the simple fact remains that Google can offer digitization on a grand scale at a price libraries can afford. When asked if Harvard was participating in any other large digitization initiatives (e.g., American Memory, Project Gutenberg, Million Book Project, OCA, Universal Library), Flecker explained, "No, we are not participating in any of these other projects at this point. Google approached Harvard with a proposal to do large-scale digitization at their expense. No one else has such made such a proposal." However, Flecker qualified, "Our agreement with Google is non-exclusive, and we would be very open to working with other digitization initiatives. Increasing the corpus of digitized materials available across the Internet is a major priority of the Harvard libraries."

Flecker's comments point to an important issue within the Google Book Search Library Project: exclusivity, or the lack thereof. As mentioned earlier, Kahle of the OCA has concerns over Google's motives to build "a private library," but in conversations with administrators at the participating libraries, non-exclusivity was a common theme. For example, when asked how digitization partnerships with current for-profit vendors such as ProQuest would be affected, NYPL's Ferriero responded unequivocally, "All agreements we sign are non-exclusive, meaning it is possible that multiple vendors could film or scan the same text." Librarians are not in the business of limiting access but rather increasing it, so partnering with Google may seem like a logical next step in any digitization program.

Moreover, partnering with Google is not the only option. In addition to digitization initiatives, such as those elsewhere in this article (i.e., American Memory, Project Gutenberg, Million Book Project, Open Content Alliance, Universal Library, Making of America), Microsoft has its own rival book-scanning project. In the Dec. 17, 2006, *Chronicle of Higher Education* The Wired Campus column, Microsoft's Live Search Books is discussed: "It may seem like Google's much-debated book-scanning project has secured the participation of every library under the sun. But Microsoft's less-discussed rival project has managed to recruit some pretty big names of its own — including the British Library, the University of California, and the University of Toronto"

[<http://chronicle.com/wiredcampus/article/1759/microsoft-releases-rival-to-googles-book-scanning-project>]. Again we see the results of librarians' commitment to non-exclusivity, which translates into a commitment to access: getting the right book to the right reader at the right time.

The public at large may be encouraged by discussions such as those within the CIC to build common digital repositories, and with not one but two participating Google Book Search libraries, the prospects look good. The simple fact that libraries are receiving their own digital copies goes a long way to allay fears of "the private library." Karle-Zenith of UM emphasized, "We are receiving our own copies of the digitized volumes so we can ensure they are preserved for future generations and made accessible as a public resource. While this may not be Google's mission, it is the mission of the library and we take this very seriously." Indeed, most libraries and librarians alike take the notion of unfettered access to library materials *very seriously*, hence the long-standing differences of opinions between librarians and publishers.

In terms of other concerns relating to participation in the Google project, all of the librarians interviewed were asked about any restrictions Google placed on their use of the digital copies (other than copyright restrictions). While the specific restrictions contained in some of the agreements are under nondisclosure, Popham of Oxford said, "We do not consider them onerous, nor an impediment to the scholarly uses to which we envisage the data might be put." Furthermore, when asked about any initial concerns relating to participation in the Google project, Popham said, "We only had two major concerns about participating in the project. Firstly, that the digitization process should not result in any more damage to the physical condition of the materials chosen, other than what we might expect to see if a reader were to consult one of the books in our reading rooms. Secondly, that we would not be unduly constrained in our ability to reuse the resulting digital data for scholarly purposes."

Karle-Zenith of UM echoes Popham's comments about initial participation reservations: "We wanted to ensure that we would receive images that adhere to library preservation standards, and we do ... We wanted a guarantee that our materials would not be damaged or destroyed in the process of

digitization ... We were concerned about having the appropriate rights to utilize our copy in ways that are consistent with the library's mission." Karle-Zenith was also able to discuss some specific restrictions: UM is required to restrict automated access to its digital copy and to take measures to prevent third parties from either downloading its copy for commercial purposes or redistributing any portions of its copy. UM must also restrict automated and systematic downloading of the image files from its copy. Indeed, at least in UM's case, these specific restrictions do not appear overly onerous, as Popham put it.

We are nowhere near to hearing the last word on how participating libraries will use the digital copies received through the Google Book Search Library Project. With more libraries climbing aboard the project on a regular basis, the possibilities are both complicated and endless. Perhaps Popham said it best: "What Google brought was an exciting vision and the resources to make that a reality."

Books with expired copyright can be downloaded directly from the site as PDFs. There is a button for that. There is a program called GoogleBooksDownloader, which is no longer supported. It allowed to download the previews Google Books allowed. As in most cases it was randomized, one could download up to 97% of a book using that program. In others it was just snippets or just the metadata, depending on deal with the author and publisher. Google has been steadily making it harder to crawl its repository over the years, though, and most old tricks don't work anymore. Specifically it is for using the search function to "search inside this book." If you found information you want to cite and reference and you get results like those pictured below Users searching with Google will see links in their search results page when there are books relevant to their query. Clicking on a title delivers a Google Print page where users can browse the full text of public domain works and brief excerpts and/or bibliographic data of copyrighted material. Library content will be displayed in keeping with copyright law. "The idea that the world's knowledge, as held through books and libraries, is opening up to all via a Web browser cannot be understated," he says. "People will find books they never knew existed." He thinks it will take a year before university and public library books begin to show up in Google's index in a meaningful way. In October, Google began working with publishers to make portions of their books searchable in the Google index. The Google Books Library Project is an effort by Google to scan and make searchable the collections of several major libraries. Along with bibliographic information, snippets of text from a book is often viewable. "Our participation in the Google Book Search Library Project will add significantly to the extensive digital resources the Libraries already deliver," said James Neal, Columbia's vice president for information services and university librarian. "It will enable the Libraries to make available more significant portions of its extraordinary archival and special collections to scholars and researchers worldwide in ways that will ultimately change the nature of scholarship."