Google Books causes concern among librarians, authors - The Boston Globe

Digital library's growth has some worried it may be building a monopoly

Dan Clancy makes librarians nervous.

When the Google Books engineering director participated in a panel discussion at the Boston Public Library this week, his opening remarks focused on the search engine's efforts to enable access for "every kid in Arkansas" to Harvard-size digital libraries. But soon afterward, he was hearing from librarians on the panel that they felt “queasy” about Google Books.

“I appreciate the trepidation, because this project is very big, and very complex,” he said afterward. "It is also part of a very large cultural and societal shift.”

Google’s growing digital book project is making some in the publishing world nervous - a fact the search giant is trying to change. Google Books, which includes the largest team of engineers working out of Google’s Cambridge
office, has been a force ever since it started an aggressive book scanning project with some of the world’s largest libraries in 2004. But now that Google has become a publishing powerhouse - with more than 10 million books scanned so far, of which 1.5 million are currently available online free of charge - it has made some librarians and authors uneasy.

“Google is creating a mega bookstore the likes of which we have never seen,” said the panel organizer Maura Marx, executive director of Open Knowledge Commons, a Boston nonprofit organization. “People are very uncomfortable with the idea that one corporation has so much power over such a large collection of knowledge.”

A growing concern, which was raised during the library panel, is that Google will end up with monopolistic control of access to millions of scanned digital books. This concern was heightened when Google negotiated a settlement with the Authors Guild and the Association of American Publishers, groups that represent authors and publishers, after they sued Google to stop the search company from digitizing books.

The complicated settlement (it runs 134 pages), which is being reviewed by Judge Denny Chin of the US District Court for the Southern District of New York, would give Google the right to display previews of books that are out of print but still in copyright, and sell access to these books to individuals and institutions. Rights holders to some of these books are difficult to locate; these books are often referred to as “orphan” books. In return, Google is paying $125 million to, among other things, establish a Book Rights Registry that will distribute proceeds to authors and rights holders from any revenue generated from these works.

Right now, the Cambridge team, which consists of 25 members, is working on creating the logistics for implementing the settlement, sorting out the rights to copyrighted publications, and adding metadata that enables more efficient searching and display of the digital books in Google’s collection. Clancy said that the number of books covered by the settlement, which has been estimated to be as high as 7 million, will be part of a much larger corpus that will continue to grow as new works are added to Google Books.

“The settlement is important, but it’s primarily about out-of-print books,” Clancy said as he sat in a small conference room in Google’s Kendall Square office. “The settlement is valuable because comprehensiveness is valuable. It would be a shame if there was this hole in the collection of online books available, because we couldn’t work out how to handle the copyrights.”

The next chapter in the Google Books project will come sometime after Oct. 7, when Chin holds a final hearing on the settlement. But Clancy indicated the Google Books team will not be standing still in the meantime.

“We’re going full steam ahead, no matter what happens with the settlement,” he said.

Indeed, the settlement is only part of Google’s plans to expand its digital book collection. The morning after the event at Boston Public Library, Clancy and Jon Orwant, who directs the Google Books operation in Cambridge, described an ongoing project that dwarfs the scope of the controversial settlement.

Google’s ultimate vision, Clancy said, is an “uber book platform” in which millions of digital books are available in a marketplace open to publishers, consumers, and online bookstores.
Clancy said that Google does not release detailed information about the number or location of its scanning centers, but that there were “around six” scanning centers around the world scanning “thousands of books a day.” At the same time, Google is building partnerships with publishers that allow Google Books to host the publishers’ content. In the Google model, digital books will be independent of any particular electronic reading device, such as Amazon’s Kindle.

“The way we see it, you should be able to buy a book for your Sony Reader, and then read it on a Barnes & Noble reader,” he said. “Consumers don’t want their books locked up in proprietary devices and walled gardens.”

This year, Google Books has already struck two deals that have telegraphed the model it is rolling out. In March, it made a half-million public domain books available to Sony’s electronic reading device. A few weeks ago, it made a similar deal with Barnes & Noble. Although the public domain books are available free, in the future the model could be married with payment systems and copyrighted books.

Clancy said the possibilities for Google Books have yet to be determined. “Will we be done when we have 40 million books? 50 million books?” Clancy said. “We don’t know yet.”

D.C. Denison can be reached at denison@globe.com.