Cycling for Libraries Berlin Seminar 6.6.2011

Discussion session 2: Why not Google?

Participants:

> Patrick / Ex Libris
>
> Andreas / Berlin University
>
> Johann / State Library of Berlin
>
> Ulrika / GBV (library network in Germany)
>
> Anssi / Espoo city library
>
> Till / GBF headquarters
>
> Lauri / Espoo city library
>
> Melanie / Free University of Berlin library
>
> Dörte / Applied University library of Schmalkalden
>
> Traian / University of Illinois
>
> Arto / CSC - The Finnish IT Center for Science
>
> Dace / National Library of Latvia

- Why is Ex Libris going through all this discovery work when Google exists
    - Why cannot we simply upload everything to Google?
      (= Why not simply open our catalogue systems and let Google crawl it?)
- There are already some strategic alliances between libraries and Google
- Users already do discovery elsewhere - what do libraries have to offer which is not found there?
    - User interfaces need to be easy: libraries do not only have a discovery problem, they have a delivery problem
- Users already do discovery elsewhere - what do libraries have to offer which is not found there?
- Awareness problem: Need to make users aware what libraries can do for them
- Why can a user find things in Google but not in a library catalogue?
- Vufind open source discovery tool (http://vufind.org/, similar to Ex Libris Primo)
    - Gathers all metadata (like Google), puts it into a single search box
- Two things: database, user interface
    - Google has both the biggest database and a friendly user interface
- Licenses create problems: cannot give open access to a lot of content
    - Often people search technical solutions for non-technical problems - the real solution would be to say to publishers that we cannot buy your stuff any more if our users cannot

use it

- What happens if library opens the library catalogue? Any experience? Some German libraries already did this. Problems:
  - Don't get the actual data if the book is available on the shelf or not
  - Those pages don't get good page ranking because page rank is based on link
  - If all libraries do the same, there's an enormous amount of duplicates
- Google suggested that the data should be available from one central server where they would harvest it
  - Already done by Austrian libraries, will be done also in Germany
- Google is good in solving technical problems - they will do that. The real question therefore is do we want that?
  - Most participants seem to agree that yes, we do
  - Discovery works on Internet, how can we get users from there to library resources
- Google is not the whole world - libraries have things to offer which Google don't have (e.g. licensed databases and journal articles)
  - Tools like Primo provide means to manage that
  - On the other hand, a small library cannot buy Primo or other expensive products - Google is seen as an affordable and easy solution
  - Aggregation and normalization of the data is a big cost - it doesn't depend whether you're using Primo or open source product such as Vuview
- What would we like Google to look like 10 years from now?
  - Search and access should be much more seamless. That's not a technical problem - it's a licensing problem. Can already be done with open access content
  - Thematic collections created by librarians and users. They can also include e.g. books which is in the library and not available electronically
- Google is not the only one - opening access to the data gives an opportunity also to other players to play with it - maybe someone will do it still better than Google?

Summary:
- It's good to have Google crawl and query your database, but it's not the only tool
- Google is a starting point but not the end
- Data must be shown in a very consistent manner, and opened
- Data discovery is not for free - even if the product is free we have to pay for work
- Many problems are not technological problems