Building a digital national library in Sweden

I will start by telling you briefly about the lines we are following in building a digital national library in Sweden:

Web harvesting

- The national library has been harvesting the web since the beginning of 1997; a search engine robot sweeps the Internet twice a year collecting Sweden-related web pages; daily newspapers are collected every day.
- Close to 300 million web pages have been archived so far.
- A special addition to the data protection law which was introduced four years ago seems to give the library the right to collect the documents and make them available on-site.
- However, this is not quite true. With respect to the Swedish copyright law the activities are still illegal; a library simply cannot download/copy documents which are not part of its collection.

Digitization

- As for digitization, our record is not that impressive. There have been scattered activities for the last five or six years: posters, rarities, weeklies, etc.; in 2006 we received extra funding from the government as a measure for generating employment; 30 people were temporarily employed to digitize out-of-copyright material, using a thematic approach. “Journeys over time” is the dominant theme.
- Only documents out-of-copyright have been made available on the web; so far we have no collective licences; the truth of the matter is that we have had great difficulties in finding contracting partners among the right holders.
- However, we are also one of the partners in a project called The Literature Bank, which is managed by the Swedish Academy. The purpose is to make the most important Swedish fiction freely accessible on the web. It is mostly out-of-copyright material that has been digitized and made available so far, but there are licences for a limited number of copyright-protected titles.

Legal deposit of electronic publications

- Sweden still does not have a legal deposit law covering web publications; however, over the last few years the National Library has signed deposit contracts, and developed functioning workflows, with a few individual publishers, mostly government agencies and universities.
A new law is expected in 2008, or 2009; of course, one of the most important questions is: how can we make this material available? One of the options considered, only on-site access at the National Library, would make it far less available than the print deposits which are distributed to all of the large universities.

**i2010 - The European Digital Library**

The European Digital Library initiative was already accurately described by Andrew Cranfield.

As I see it, the positive sides are, i.e.,

- That it underlines the importance of preserving the European cultural heritage and making it available, thereby providing moral support to heritage institutions trying to squeeze out funding for digital library projects.
- That it supports cross-border knowledge development and networking among archives, libraries, and museums.
- That it supports the development of digitization centres designed to meet the needs of the heritage institutions in cost-effective ways. The option of doing all digitization in-house is not realistic for most national libraries.
- That it supports the development of best practices based on international standards.

But there are problems:

- No financial support is given to actual digitization.
- Restrictive copyright regulation, for which the EC is highly responsible, seriously limits the number of publications which can be digitized and made available - this in turn makes it more difficult to attract funding nationally, be it from public or private sources.
- We are going to have problems making archives and museums accept the leading role of CENL - and on the other hand, the CENL also seems to have some problems with accepting the participation of archives and museums.
- Integrating the digital collections of heritage institutions from close to 50 European countries, e.g. coping with multilinguality, is also a formidable task.
- Obvious overlapping with similar projects, like the museum-driven project MICHAEL and the French Europeana creates a lot of confusion; and so does overlapping between different CENL projects.

Are there any solutions to these problems?

- EC funding for digitization? I do not think it is too far-fetched. The Commission is indeed subsidizing ongoing, regular activities in other sectors. BUT, I don't think it is going to happen.
- Changes in copyright regulation? The Commission got us into a mess concerning copyright. Why should it not get us out of it, for instance by shortening the term of protection? Here, I don't know if there is any hope. We must remember the copyright system is a global one, and
any changes in a more liberal direction would take considerable time. We must also remember
that in the digital environment copyright regulation is more and more dealt with in contracts
and licences.

Licences? This is a solution recommended by the commission. And licences are here already.
But mostly they cost money, at worst a lot of money, and then it’s doubtful whether you can
talk about “solutions”. My hope is that the licences will be priced according to the commercial
value of the publications concerned, or why not the actual use, because that would make it
pretty cheap, but we already see proof of the opposite, i.e. gross over-pricing.

Merging EDL and MICHAEL has been suggested as one way of diminishing overlapping. That
seems to be a great idea.

My vision of the European Digital Library:

I would like to see the European Digital Library as the sum of the digital European heritage
collections: text, sound, images. Irrespective of what portals or what search engines are used
to access them.

I’m not at all opposed to the construction of a portal called the European Digital Library, but
these collections should be exposed in many different ways and contexts (by local web sites,
national web sites, European web sites, global web sites, and by multiple search engines).
Maximum exposure is what we want. Multi-purposing is the concept. OAI-PMH is one means of
achieving it. This way, we will not be totally dependant on the success or failure of one project
or other supported by the EC.

Some final remarks:

The idea of the European Digital Library, and that of the Information Society 2010, partly goes
back to a perception that the Union should compete with the other economic super powers of
the world in the area of information provision. Commissioner Bangemann talked about this as
early as 1985, and these arguments are still being repeated now and again.

I am definitely positive to the EC’s initiative, but I can see some risks in focusing too exclusively
on the European perspective. Do we have to stick to the European agenda EVEN IF it should
counteract a more global approach? EVEN IF it should disregard the option of partnerships with
non-European players, whether they’re commercial giants like Google and Microsoft, or non-
commercial players like, e.g., the Open Content Alliance. EVEN IF it should mean going against
national priorities? Or, EVEN IF it should mean getting stuck in a bureaucratic structure that
produces very little in terms of concrete results?

I think we will end up on the wrong track if we keep saying things like “we will challenge
Google”. And we should remember that in all the time, from the beginning of this millennium
onwards, that the EC has supported digitization, the member countries have altogether
produced far less digitized documents than South Korea alone.

Of course there is no definite answer to my questions, and admittedly they are still hypothetic,
but I don’t think they’re entirely irrelevant.