68th IFLA Council and General Conference
August 18-24, 2002

A multifaceted strategy for a National Bibliography on electronic resources

Randi Diget Hansen
The Danish Bibliographic Centre
Copenhagen, Denmark

Abstract

The Process of creating a National strategy in Denmark concerning registration of net-publications.

The National strategy is concerned with the following efforts:
Clarifying the quantity of Danish netpublications candidating for registration in the national Bibliography.- Construction of a database based on Dublin Core where all Danish produced net-registrations can be stored.- Construction of an intelligent agent as an automatic help to the cataloguers.- Contracting strategic alliances with producers, libraries and other relevant institutions.- Experimenting with various levels of registration and different registration formats.
The conclusion is that the amount of net-publications to be included in the National bibliography is not overwhelming and that bibliographic control of the internet is manageable with re-use and co-operation.

The situation in Denmark

The proces of creating a National Strategy in Denmark concerning registration of net-publications.
The strategy contains these different purposes:

-   Creating a model for co-operation between different types of Internet services in the library world and
    The National Bibliograhic Agency – including design of a database for sharing of resources, creation

of exchange format, development of selection criteria and other kinds of tools and rules that can facilitate the work.

- Convincing the contributing authorities and the people in the library world as a whole that bibliograhic control of the Internet is necessary and practical and that re-use and co-operation is the only way of making it feasible.

National bibliograhy and legal deposit in Denmark concerning net-publications

In 1996 the Danish Bibliographic Centre, producing the Danish National Bibliograhy, decided to launch a project on finding out whether net-publications could or should be subject to bibliograhic control in the same way as printed and electronic publications in fixed physical form. The reason for this was that we felt the existing search engines on the net suffered from the general problem of searching in unqualified data and generally replying with excessive amounts of data. We also felt that the information contained in net-publications not necessarily differed from the information in publications in fixed physical form, and if Internet publications are excluded from bibliographic control there is the risk that many people will find it difficult to gain access to an increasing amount
of the information citizens' need as opposed to the information stored in products in fixed physical form.

To the question whether it is worth registering publications that may disappear again we concluded that it is important to focus on the "quality" of the publication registered and we felt convinced that net-publications will tend to become more permanent.

Simultaneously a new law of legal deposit was prepared and took effect from 1st January 1998. This law contains all information carrying media including publications of digital works in databases – though restricted to static publications i.e publications, which look like or replace books.

The new law of legal deposit implied that Denmark also since 1998 has had a National Bibliography in the area covering static net-publications.

From the beginning we have had the assumption that we are not talking about millions and millions of publications to be registered but that the amount of qualified publications on the Internet which can meet the inclusion criteria of the National Bibliography is not bigger than the task is practical if it is based on co-operation with producers and other networkers in the library community.

We have also concluded that the idea of publications being transient and will disappear – we all know the calculation that the average life of a document on the Internet is 42 days – is not valid when we talk about qualified publications.

Finally, we have been convinced that the great task of organizing qualified information on the Internet demands a constant process of development. It is not realistic to believe that the library community in no more than 6 - 7 years have found the final models for handling net-publications.

What has happened since 1998 - why do we work with a national strategy?

The law of legal deposit and The National Bibliography are still restricted to static publications. On the other hand there is a lot of registration going on in the library community. The contributing authorities have been willing enough to pay millions of Danish crowns to initiatives that could be defined as projects, but not to encrease the national bibliography so dynamic publications could be included.

There are lots of projects as well in the public library sector as in the sector of State and University libraries. There are portals for children, music, food, health services etc. but also several general portals.

The Danish Bibliographic Centre has produced about 12.000 registrations consisting of Danish static and dynamic publications and Danish articles and reviews. The other net-services in the library sector that are subsidiced from the Danish State have produced about 20.000 registrations of relevant resources in any languages all with human interference. Unfortunately, despite the size of Denmark, the different projects use various technical formats and also extensive amounts of double registrations take place.

The reason why the Danish State until now has not wanted to pay for an extension of the National Bibliography is the fear of exploding amounts.

The National strategy is concerned with the following efforts

- The " Statistics" project where the aim is to clarify the quantity of Danish net-publications candidating for registration in the National Bibliography

- Construction of a database "The Metanet " based on Dublin Core where all Danish produced net-registrations can be stored

- Construction of an intelligent agent " The Netowl" as an automatic help to the cataloguers in the collection process. It is trained like a child to find out what is good and what is not wanted.

- Strategic alliances with producers, libraries and other relevant institutions.

- Experimenting with various levels of registration and different registration formats.

The Statistics project

DBC has in the autumn of 2001 and the spring of 2002 carried through two pilot projects to clarify the quantity of Danish net-publications candidating for registration, and both of them have confirmed our assumption that the amount of publications is managable.

The project is based on a random test of 2000 subdomaines of the in all 351.792 .dk subdomaines. We only concentrated on the .dk domaines.

Of the 2000 subdomaines 1469 could be found but only 732 had contents. The others were empty, reserved, had an automatic reforwarding, were double URLs etc. Of the 732 domaines with contents only 46 had relevant contents according to the inclusion criterias of the National Bibliography. In all, 76 relevant publications were found.

If the tendency is tenable – and it has been confirmed in two testings – it means that for the time being there are in the .dk area about 14.000 publications worthy of inclusion in the Danish National Bibliography and an estimatet increase per year of about 4.000 publications. Since 2000 the tendency of increase in new domaines has been slightly falling.

The "Metanet" database

The Metanet database which is based on the Dublin Core format is established at DBC. It is thought as a joint pool for all Danish registrations. The suppliers are DBC - as the central supplier of bibliographic data; this counts for both national bibliographic records and library cataloguing - as all Danish libraries working with net-registrations both Public and Research libraries. It is important to profiling own products therefore a lot of double, tripple-registrations are going on, which is bad utilization of the amount available. Theaim is re-use and rational registration of the Internet. It has also been necessary to define an exchange format based on DC with administration components which are coping with all the different technical formats used in the net-services. You can find it on DBC's homepage.

The " Netowl"

One of the great challenges if not the greatest is still how to collecting and keeping the net-registrations worthy of being a part of the National Bibliography under surveillance in a methodical way. Some are delivered from the legal deposit administration www.pligtaflevering.dk but we know it is not yet a well-established routine for alle producers of net-publications to notify their publications to the legal deposit registration form. And it is still only static publications which are part of the legal deposit. The "Netowl" is developed as an automated help in this surveillance work. It is a search engine trained by the cataloguers. It is based on boolean algebra. The training has just started and people are in the optimistic phase.

Strategic alliances

Another tool to ensure a more methodical and covering collection of net-publications is to make alliances with producers, libraries and institutions which have an interest in providing net-publications for their users in the catalogue and want to ensure the durability of the resource. As an example The Administrative Library in Denmark has an obligation to provide and store all publications coming from ministeries, government departments and agencies. This Spring they have just received grants to storing the publications forever. They want to give the permanent URL to the National Bibliography and they want to make a methodical campaigne to collect the publications and the National Bibligraphic Agency will register them. The project will start up this september.

Experiment with various levels of registration and different formats

In Denmark it is an obligation to use the Danish Cataloguing rules and the danMarc2 format in the National Bibliography. It is described in The National Bibliography Agreement. It is also valid for net-publications. So this is also the rules and format we used when starting up the national bibliography in this area. Those are good rules and a good format. They may, however, make you think in a very traditional way, and people escpecially non cataloguers think it is very time consuming and expensive. We have therefore used the freedom in a new type of material to experiment. We have engaged librarians without any training in cataloguing and made them use a minimum level in Dublin Core. It seems to be efficient and with a satisfying product. We think people should concentrate on the difficulties of selection and subject treatment not description.

Why a national strategy for registration of net-publications

Why do we in Denmark think that a national bibliography covering net-publications is necessary and practical.

The National Bibliography is the memory of a nation. It is a task to secure that published knowledge and experience produced in Denmark is secured and registered for actual use but also for the future. This is an obligation in whatever form the national imprint has.

The publications produced by Danish authors, artists, firms, governmental organizations etc. which are only published on the Internet will only be secured in a methodical way if the work is done or co-ordinated by an agency like the National Bibliographic Agency which has the task of methodical, continuous registration.

You cannot rely on the fact that even the best search engines will make a methodical job and take on the obligation of archiving and making available. The purpose of the search engine is another, namely giving admission to the Internet pages existing at whatever given moment.

The " Statistical " project has proved that the ideas of the overwhelming job it is to make a kind of manual registration of quality publications is not real. It seems as if the amount of publications to be included in the national bibliography is manageable.

The stability of the registered publications also seems to be acceptable. In 2001 DBC made an URL check of the 4.000 registrations in the National Bibliography and we only found 20 permanently disappeared publications. This means that the publications we select are stabile. The co-operation with The Administrative Library means that from now on all State publications are secured from disappearing. A great number of static publications are delivered for legal deposit but unfortunately due to statutory reasons they are not available. The development should be that both static and dynamic net-publications selected for registration in the National Bibliography as a minimum are delivered for legal deposit and stored. It should be the existing cultural institutions' interest to carry out this task.