

Electronic publications – access now and in the future

A seminar at the Royal Library, Stockholm, Sweden,

18 October 2005

Tomas Lundén, Librarian
Lund University Libraries, Head Office

Published in: *ScieCom Info* 2005:3 (November 2005)
ISSN: 1652-3202

Alternative location:

http://www.sciecom.org/sciecominfo/artiklar/lunden_05_03.shtml

Electronic publications – access now and in the future

A seminar at the Royal Library, Stockholm, Sweden,

18 October 2005

Tomas Lundén, Librarian
Lund University Libraries, Head Office

Introduction

Swedish universities and government authorities are increasingly making their publications electronically available on the World Wide Web, either on web sites or in local systems, so called institutional repositories. Standardisation of metadata is a way of increasing access to these publications. Tools and workflows are also being developed to secure long term preservation and access. On 18 October 2005, a seminar presenting and discussing different initiatives and issues in this area, was held at the Swedish Royal Library (KB). The seminar was arranged by BIBSAM, the Royal Library's Department for National Co-ordination and Development. Entitled "Electronic publications – access now and in the future" the seminar was broad in scope, ranging from technical and organisational issues to political ones, taking in perspectives from both universities (and their libraries) and government authorities, as well as from KB. On the agenda was also the matter of legal deposits of electronic material.¹

In this brief report, I will focus on the issues raised that concern scientific publications, mainly the results from the SVEP project and the question of legal deposits of electronic material.

The SVEP project

The SVEP project was presented in brief by project coordinator Jan Hagerlid (who was also the moderator of the seminar). SVEP aims at promoting, coordinating and supporting electronic publishing at Swedish universities and university colleges. The project started in September 2003 and is finished by Autumn 2005.

The project has been funded and coordinated by BIBSAM and carried out in close collaboration between nine Swedish university and university college libraries and KB. The major participants besides KB is Uppsala University Library and Lund University Libraries. The project was divided into five work packages, two of which were presented at this seminar.²

Stefan Andersson from Uppsala University Library, head of WP 1, presented the work concerning interoperability. The project has produced recommendations for metadata (description) formats for electronically published scientific publications, and also common subject categories. The subject categories are based on categories on Swedish research used by Statistics Sweden (Statistiska Centralbyrån) and the Swedish Research Council. These are recommendations that are intended to serve as a guide for universities and university colleges that maintain and develop (or are in the process of starting to

develop) institutional repositories for electronic fulltext publications. The implementation of the recommendations will promote the exchange of information and is a prerequisite of developing advanced search services.

The recommendations for fulltext documents function on two levels. On the minimal level the documents will be made available for international service providers through OAI-PMH³ and on the expanded level (“SVEP level”) they will form the basis for advanced search services and be compatible with other bibliographic databases, for instance Libris, the national library catalogue.

National recommendations have also been developed for local research databases (where the full academic output of a university is registered as references, not necessarily with fulltext linking). The adoption and implementation of these recommendations by universities and research institutions would make it possible to for example conduct comparative studies between universities and maintain statistics. It could also be desirable to make Swedish research visible through a joint search service, based on these recommendations. The board of the Association of Swedish Higher Education (SUHF) decided on 31 August to adopt the recommendations.

So all in all, there is a lot to be gained through common formats and standards, perhaps most importantly to avoid duplication of efforts, to make the process of registration and publication as automated as possible, and through intelligent use and re-use of data and metadata support maximum visibility and accessibility of Swedish research output.

Another very important issue concerns long term preservation of electronic publications. This has been the object of WP 2 in SVEP. The aim has been to create a workflow and technical solutions for long term preservation of electronically published documents at Swedish universities.

Considering the development pace of software there is today no guarantee that a specific document or file format will be rightly interpreted in the future, or possible to open and read. As Eva Müller, head of WP2, from Uppsala University Library made clear, these issues are important from several different viewpoints: for researchers (and students), it is of course important to know that their work will be possible to find, read and interpret in the future. For universities, it is important for keeping track of their own research results. For libraries, it is of course part of their fundamental task of providing access to information. And in a larger perspective, it is a matter of preserving our cultural heritage for future generations.

WP 2 has developed the groundwork for an infrastructure which will make easier the process of gathering and preserving documents and their metadata. This is meant to generate a standards-based, automated workflow between the local repository and the Royal Library (KB). The idea is a system of persistent identification of documents, called URN:NBN.⁴ A resolution service will keep track of the document and point to the current URL where the document is located. If the URL changes the resolution service will point to the new address thanks to the persistent identifier. Since a copy of the locally published document is to be kept at KB, even if the local repository closes down and its data is lost forever, the resolution service will still point to KB’s copy.

To automate the workflow between the local system and KB, WP2 has been working on something called standardised information packages. These packages contain the document and appropriate metadata (including technical and rights metadata). A prototype for a package tool has also been developed, whereby the creation of the packages could be standardised.

Another part of the workflow is the development and maintenance of a format registry, where information on software and file formats will be stored. This of course will be vital information for securing future access to electronic documents.

The format registry and the package tool are still only prototypes, and the workflow is yet to be realised in practice. This is however the object of an upcoming project lead by Uppsala University Library, to further develop tools and establish the workflow. No doubt an important part in establishing this will be played by KB.

Legal deposit of electronic material?

KB's role in the question of long term preservation is to a large extent dependant on if or when a law regulating the legal deposits of electronic material (e-deposits) will be made. There is currently the Statutory Deposit of Copies Act (SFS 1993:1392) regarding printed material (and Sweden has had legal deposit of copies of printed material since 1661), and the e-deposit report was published in 1998 (SOU 1998:111), but has not so far led to any legislation.

Gunilla Jonsson from KB brought this up in her presentation, which dealt with a project on digital deposit to KB. KB has had digital deposits since 2000, and the current project has aimed to improve the technical solutions and develop an infrastructure for deposit of electronic material. The project was similar to SVEP's WP2 and part of the workflow is a collaboration between the two projects. (It would seem to be a good idea for KB to further coordinate this work in the future.)

The digital deposits originally grew out of an interest from the external parties (small webbased publishers, universities, government authorities etc.) that wanted to preserve their digital output but did not know how or did not have the finances for it. The deposit of material is wholly optional, but is bound by an agreement between KB and the supplier. KB has also since 1997 an automatic harvesting of Swedish web pages, the Kulturarw3 project. What is needed now are guidelines for what the suppliers should deliver and what material can be left to the automated Kulturarw3.

With the support of an e-deposit legislation, Jonsson argued, KB could more efficiently build on their experiences and develop further automated functions in the workflow.

Susanna Broms, Legal Advisor at KB, talked about e-deposits from the legal perspective. She pointed out that several countries have varying kinds of legislation on e-deposits, among them Norway, Denmark, France and Great Britain. The major obstacle in the current system of harvesting and depositing e-material is caution of infringement of copyright and the Personal Data Act (Personuppgiftslag, SFS 1998:204). There is a proposal for a law on e-deposits, this would put the responsibility of delivering the material on the provider and that would mean avoiding the copyright restrictions. Regarding the Personal Data Act, this was stressed by Maria Ljungkvist from the Division for Research Policy, an operational division under the Ministry of Education,

Research and Culture. The act aims to prevent the violation of personal integrity in the processing of personal data, and the Ministry of Justice has according to Ljungkvist, tended to regard even reference lists in official reports and documents as possible sources of violation, if processed in the wrong way.

However, the government wants to investigate the question of e-deposit, as is said in the latest Government research policy bill (2004/05:80).⁵ Ljungkvist argued that a new act on e-deposit will be proposed some time in 2006 at the earliest. And since 2006 is an election year, we cannot know how a new government would prioritise the issue. So when an e-deposit legislation actually will come is still unclear.

During the presentations on e-deposits, the issue was raised on what type of material that should be encompassed by an e-deposit legislation. And of course there is a huge diversity of material on the web, so definitions of this is much needed. For example, personal web pages could perhaps be included in Kulturarw3, but not necessarily included in legal deposits. E-deposits should encompass documents with durability, rather than perhaps, dynamic web pages.

At the same time, if this is the case, then we must consider the changing world of scientific publication. Increasingly, research articles can be of a dynamic character, contain multiple objects in different file formats, streaming content, links to data sets and databases etc. So definitions will have to take into account not only technical format, but content and purpose of publication as well.

Discussion and final remarks

Gunnar Sahlin, National Librarian, stated in his concluding remarks that the seminar had focused on questions of the utmost importance, electronic publications and their preservation for the future. The development of electronic journals and other electronic media has of course been vital for the research communities. And the infrastructural matters discussed and presented are essential for the research itself to function. He pointed out that the European Union has acknowledged the need for digital preservation, and that it is included in EU's Seventh Framework Programme, and hopefully this will lead to a coordination of efforts within the Union.

An interesting question was asked regarding coordination of registration of Swedish research output. Instead of relying on data from Thomson ISI (since ISI has limits), could a national initiative for gathering this data be started? Jan Hagerlid answered that the question has been investigated for some years, and that the Swedish Research Council was interested in building its own database. It was decided that this would probably violate copyrights, since the data would be based on ISI's data. But hopefully the recommendations that was produced in SVEP will be adopted and implemented by universities in Sweden, which would make it possible to use the data as a basis for statistics.

Sahlin concluded that better coordination is the next step. KB has more frequent contact with both the Swedish Research Council (VR) and the Association of Swedish Higher

Education (SUHF), and the organisations have a greater understanding that cooperation in matters of research infrastructure is needed.

Svensk sammanfattning

Svenska högskolor och myndigheter publicerar alltmer material elektroniskt på webben. Den 18 oktober 2005 hölls ett BIBSAM-arrangerat seminarium på KB, betitlat ”Elektroniska publikationer - tillgänglighet nu och i framtiden”. Det var ett brett seminarium som behandlade olika aspekter av e-publicering, tekniska, organisatoriska och politiska. Standardisering av metadata, verktyg och arbetsflöden för att säkra tillgång och bevarande i framtiden, var exempelvis saker som togs upp. Denna artikel berör frågor som främst rör vetenskapliga publikationer.

SVEP-projektets resultat presenterades, med fokus på två delprojekt, DP1 och DP2. DP1: Interoperabilitet, presenterat av projektledaren Stefan Andersson, har utvecklat rekommendationer för metadatabeskrivningar av elektroniskt publicerade vetenskapliga publikationer, samt även för s.k. publiceringsdatabaser (lokala register för akademisk publicering). Syftet är att genom en harmonisering av metadatan främja utbyte av information och ge förutsättningar för olika tjänster. DP2: Långtidsbevarande, som presenterades av Eva Müller, har utarbetat grunden till en infrastruktur för att kunna bevara och återfinna elektroniska dokument i framtiden. Systemet bygger på ett automatiserat arbetsflöde mellan det lokala institutionella arkivet och KB, där flera olika komponenter ingår, som en uppslagstjänst och ett format- och datamiljölexikon. Dessa är ännu prototyper, men ska utvecklas vidare.

På seminariet togs också frågan om e-plikttag upp. Gunilla Jonsson tryckte på vikten av en sådan lag för KB:s del, för att man effektivt ska kunna fortsätta utveckla automatiserade funktioner och förbättra hanteringen kring digitala leveranser. Susanna Broms tog upp det juridiska perspektivet, och presenterade det förslag om e-plikttag som finns. Upphovsrätten och personuppgiftslagen är frågor som är knutna till insamlande av elektroniskt material och som Maria Ljungvist från Forskningspolitiska enheten framhöll, är det frågor som måste lösas innan e-plikttag är ett faktum. Men möjligen kan det bli en proposition under 2006. Beroende bl.a. på resultatet i riksdagsvalet hösten 2006, så kan man dock inte säga exakt när en e-plikttag kommer.

Riksbibliotekarie Gunnar Sahlin avslutade seminariet med att poängtera att mer samordning kring forskningens infrastrukturella frågor är nästa steg. KB har tätare kontakter med Vetenskapsrådet och SUHF och en större förståelse finns för samarbete i dessa frågor.

¹ The full programme (in Swedish) with powerpoint presentations is available at <http://www.kb.se/BIBSAM/kursokonf/ovriga/elpubl18okt2005.htm>

² The results of the other work packages: WP3 has built a national search service for Swedish undergraduate theses and diploma work, while WP4 and WP5 has worked with advice and support on e-publishing – tools, standards, overview of software, seminars and workshops. Full information (in Swedish) including final reports and an independent evaluation of SVEP can be found on the project's web site: <http://www.svep-projekt.se/>

³ Open Archives Initiative - Protocol for Metadata Harvesting, is a standard whereby electronic publications in a local repository can be harvested in a simple format by service providers globally. See: <http://www.openarchives.org/>

⁴ Uniform Resource Name:National Bibliographic Number. More info: <http://www.kb.se/urn/>

⁵ Government research policy bill 2004/05:80 (Forskningspolitiska propositionen), pp. 108-111. <http://www.regeringen.se/content/1/c6/04/11/35/6effb2fa.pdf>