erpastudies

swedish university of
agricultural sciences

www.erpanet.org

ELECTRONIC RESOURCE PRESERVATION AND ACCESS NETWORK

Table of Contents

Executive Summary

The Swedish University of Agricultural Sciences (SLU) is very active in the field of scientific research. The SLU archives, research and IT departments realise that scientific material has potential value for future researchers, but only if it remains accessible and readable. While efforts to date have been slowed by a lack of a proper organisation and senior management support for policies for digital preservation, the SLU is now involved in a joint project with thirteen other Swedish universities to research the digital preservation of digital scientific data. This pilot project is currently in phase one and will be evaluated in late October 2004. The results of this pilot project could be of great benefit to the wider scientific and academic communities.

Chapter 1: The ERPANET Project

The European Commission and Swiss Confederation funded ERPANET Project[1] (Electronic Resource Preservation and Access Network) works to enhance the preservation of cultural and scientific digital objects through raising awareness, providing access to experience, sharing policies and strategies, and improving practices. To achieve these goals ERPANET is building an active community of members and actors, bringing together memory organisations (museums, libraries and archives), ICT and software industry, research institutions, government organisations, entertainment and creative industries, and commercial sectors. ERPANET constructs authoritative information resources on state-of-the-art developments in digital preservation, promotes training, and provides advice and tools.

ERPANET consists of four partners and is directed by a management committee, namely Seamus Ross (HATII, University of Glasgow; principal director), Niklaus Bütikofer (Schweizerisches Bundesarchiv), Hans Hofman (Nationaal Archief/National Archives of the Netherlands), and Maria Guercio (ISTBAL, University of Urbino). At each of these nodes a content editor supports their work, and Peter McKinney serves as a co-coordinator to the project. An Advisory Committee with experts from various organisations, institutions, and companies from all over Europe give advice and support to ERPANET.

---

[1] ERPANET is a European Commission funded project (IST-2001-32706). See www.erpanet.org for more details and available products.

Chapter 2: Scope of the Case Studies

While theoretical discussions on best practice call for urgent action to ensure the survival of digital information, it is organisations and institutions that are leading the drive to establish effective digital preservation strategies. In order to understand the processes these organisations are undertaking, ERPANET is conducting a series of case studies in the area of digital preservation. In total, sixty case studies, each of varying size, will investigate awareness, strategies, and technologies used in an array of organisations. The resulting corpus should make a substantial contribution to our knowledge of practice in digital preservation, and form the foundation for theory building and the development of methodological tools. The value of these case studies will come not only from the breadth of companies and institutions included, but also through the depth at which they will explore the issues.

ERPANET is deliberately and systematically approaching disparate companies and institutions from industry and business to facilitate discussion in areas that have traditionally been unconnected. With these case studies ERPANET will broaden the scope and understanding of digital preservation through research and discussion. The case studies will be published to improve the approaches and solutions being developed and to reduce the redundancy of effort. The interviews are identifying current practice not only in-depth within specific sectors, but also cross-sectorally: what can the publishing sector learn from the aeronautical sector? Eventually we aim to use this comparative data to produce intra-sectoral overviews.

This cross-sectoral fertilisation is a main focus of ERPANET as laid out in its Digital Preservation Charter.[2] It is of primary importance that disparate groups are given a mechanism through which to come together as best practices for digital preservation are established in each sector.

Aims

The principal aims of the study are to:

- build a picture of methods and match against context to produce best practices;
- accumulate and make accessible information about practices;
- identify issues for further research;
- enable cross-sectoral practice comparisons;
- enable the development of assessment tools;
- create material for training seminars and workshops; and,
- develop contacts.

Potential sectors have been selected to represent a wide scope of information production and digital preservation activity. Each sector may present a unique perspective on digital preservation. Organisational and sectoral requirements, awareness of digital preservation, resources available, and the nature of the digital object created place unique and specific demands on organisations. Each of the case studies is being balanced to ensure a range of institutional types, sizes, and locations.

---

[2] The Charter is ERPANET's statement on the principles of digital preservation. It has been drafted in order to achieve a concerted and co-ordinated effort in the area of digital preservation by all organisations and individuals that have an interest and share these concerns. http://www.erpanet.org/charter.php.

The main areas of investigation included:

- perception and awareness of risk associated with information loss;

- understanding how digital preservation affects the organisation;

- identifying what actions have been taken to prevent data loss;

- the process of monitoring actions; and,

- mechanisms for determining future requirements.

Within each section, the questions were designed to bring organisational perceptions and practices into focus. Questions were aimed at understanding impressions held on digital preservation and the impact that it has had on the respective organisation, exploring the awareness in the sector of the issues and the importance that it was accorded, and how it affected organisational thinking. The participants were asked to describe, what in their views, were the main problems associated with digital preservation and what value information actually had in the sector. Through this the reasons for preserving information as well as the risks associated with not preserving it became clear.

The core of the questionnaire focused on the actions taken at corporate level and sectoral levels in order to uncover policies, strategies, and standards currently employed to tackle digital preservation concerns, including selection, preservation techniques, storage, access, and costs. Questions allowed participants to explore the future commitment from their organisation and sector to digital preservation activities, and where possible to relate their existing or planned activities to those being conducted in other organisations with which they might be familiar.

Three people within each organisation are targeted for each study. In reality this proved to be problematic. Even when organisations are identified and interviews timetabled, targets often withdrew just before we began the interview process. Some withdrew after seeing the data collection instrument, due in part to the time/effort involved, and others (we suspect) dropped out because they realised that the expertise was not available within their organisation to answer the questions. The perception of risks that might arise through contributing to these studies worried some organisations, particularly those from sectors where competitive advantage is imperative, or liability and litigation issues especially worrying. Non-disclosure agreements that stipulated that we would neither name an organisation nor disclose any information that would enable readers to identify them were used to reduce risks associated with contributing to this study. In some cases the risk was still deemed too great and organisations withdrew.

Chapter 3: Method of Working

Initial desk-based sectoral analysis provides ERPANET researchers with essential background knowledge. They then conduct the primary research by interview. In developing the interview instrument, the project directors and editors reviewed other projects that had used interviews to accumulate evidence on issues related to digital preservation. Among these the methodologies used in the Pittsburgh Project and InterPARES I for target selection and data collection were given special attention. The Pittsburgh approach was considered too narrow a focus and provided insufficient breadth to enable full sectoral comparisons. On the other hand, the InterPARES I data collection methodology proved much too detailed and lengthy, which we felt might become an obstacle at the point of interpretation of the data. Moreover, it focused closely on recordkeeping systems within organisations.

The ERPANET interview instrument takes account of the strengths and weaknesses from both, developing a more focused questionnaire designed to be targeted at a range of strategic points in the organisations under examination. The instrument[3] was created to explore three main areas of enquiry within an organisation: awareness of digital preservation and the issues surrounding it; digital preservation strategies (both in planning and in practice); and future requirements within the organisation for this field. Within these three themes, distinct layers of questions elicit a detailed discovery of the state of the entire digital preservation process within participants' institutions. Drawing on the experience that the partners of ERPANET have in this method of research, another important detail has been introduced. Within organisations, three categories of employee were identified for interview: an Information Systems or Technology Manager, Business Manager, and Archivist / Records Manager. In practice, this usually involved two members of staff with knowledge of the organisation's digital preservation activities, and a high level manager who provided an overview of business and organisational issues. This methodology has allowed us to discover the extent of knowledge and practice in organisations, to understand the roles of responsibility and problem ownership, and to appreciate where the drive towards digital preservation is initiated within organisations.

The task of selecting the sectors for the case studies and of identifying the respective companies to be studied is incumbent upon the management board. They compiled a first list of sectors at the very beginning of the project. But sector and company selection is an ongoing process, and the list is regularly updated and complemented. The Directors are assisted in this task by an advisory committee.[4]

---

[3] See http://www.erpanet.org/studies/index.php. We have posted the questionnaire to encourage comment and in the hope that other groups conducting similar research can use the ideas contained within it to foster comparability between different studies.

[4] See www.erpanet.org for the composition of this committee.

Chapter 4: Swedish University of Agricultural Sciences (SLU)

Sveriges lantbruksuniversitet – Swedish University of Agricultural Sciences (SLU) was established in 1977. The main goal of the SLU is to foster the development of learning and expertise in the life sciences. Subjects of research expertise at SLU include agriculture, forestry, veterinary medicine, biotechnology, environmental sciences and food sciences. Currently, there are more than 5000 undergraduate and 800 postgraduate students enrolled at SLU. The SLU has main campuses at Alnarp, Skara, Ultuna, and Umeå, but research and teaching activities are carried out throughout Sweden.

The four faculties at SLU are Forest Sciences; Veterinary Medicine and Animal Science; Natural Resources and Agricultural Sciences; and Landscape Planning, Horticulture and Agricultural Science. Together, they comprise over fifty academic departments and centres. A total of 3200 people are employed at the University. The SLU Archives, Finance Office and IT Department are all located within the administrative section of the overall university structure. The SLU Archives are located within the Central Office of Administration, the Finance Office is located within the Office of Strategy and Planning and IT is located within the IT Office.

Below is a graphic representation of the SLU structure.[5]
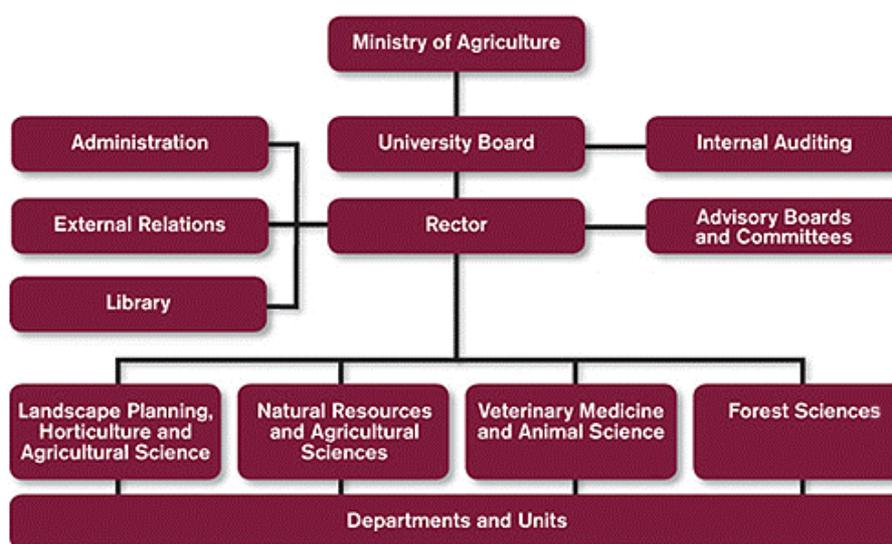
SLU is chiefly concerned with the preservation of their administrative and educational records and their scientific research output. These resources may be produced in a wide range of formats – from databases to emails to electronic publications.

---

[5] SLU Organisational Structure: http://www.slu.se/eng/thisisslu/presentation/index.html.

Chapter 5: Details and circumstances of the Interviews

Initial contact was made with Renata Arovelius, Head of SLU Archives, during the ERPANET seminar on the Selection, Appraisal and Retention of Digital Scientific Data in December 2003. Subsequent contact was made following the event and Ms. Arovelius contacted the appropriate staff in the IT and financial departments at her institution. Carl-Erik Lundgren, Chief of the IT Office, and Lars Torell, Controller Function of the Financial Office, contributed to the completion of the questionnaire. Email communication continued throughout May 2004 for further information regarding responses.

Organisation Details:

SLU (Swedish University of Agricultural Sciences)
P.O. Box 7070
SE-750 07 Uppsala
Sweden
Tel: +46 (0)18-67 10 00
Fax: +46 (0)18-67 20 00
E-mail: registrator@slu.se

http://www.slu.se/
http://www.slu.se/eng/index.html

Chapter 6: Analysis

This section presents an analysis of the data collected during the case study. It is organised to mirror the sequence of topics in the questionnaire.

- Perception and Awareness of Digital Preservation

- Preservation Activity

- Compliance Monitoring

- Digital Preservation Costs

- Future Outlook

Perception and Awareness of Digital Preservation

Perception of the importance of digital preservation appears to vary throughout the university. Participants indicate that awareness is high among the scientific departments in the university, especially those working in the fields of forestry and those working on long-term projects. Awareness is also very high within the university archives and the IT department. However, a lack of sufficient awareness and coordination from senior management has prevented large-scale, internal attempts to tackle digital preservation to date. This is an on-going problem, but the IT department and SLU archives are currently in the process of drafting a new IT policy to incorporate long-term digital preservation elements and hope that this will help to inform senior management of the potential benefits.

Information about digital preservation is gained through communications with the archive and library communities, government agencies and through attendance at digital preservation workshops and seminars, including for example the ERPANET Seminar on the Selection, Appraisal and Retention of Digital Scientific Data.[6] The SLU archives also cited web-based digital preservation advice from ERPANET and the Digital Preservation Coalition (DPC)[7] as a valuable reference resource.

The Main Problems

In practical terms, the main problem for the SLU lies in securing funding for the long-term preservation of their digital resources. On a more philosophical note, SLU archive staff indicates the lack of allocation of responsibility for digital preservation within the university and the lack of an organisational policy supported by the university's senior management as obstacles impeding progress with preserving the university's digital scientific material. The SLU archives hope that their involvement in a joint pilot project with thirteen other Swedish universities will help to highlight the potential benefits of digital preservation activities to senior management.

Asset Value and Risk Exposure

While no formal risk analysis has been conducted by SLU, they indicate that the main reason for preserving their administrative, research and education records is their legal

---

[6] The ERPANET Seminar on the Selection, Appraisal and Retention of Digital Scientific Data was held at the Biblioteca Nacional in Lisbon from December 15-17, 2003. To view the final report for the seminar, see http://www.erpanet.org/events/2004/lisbon/.
[7] Preservation Management of Digital Materials – A Handbook by Maggie Jones and Neil Beagrie. http://www.dpconline.org/graphics/handbook/.

and historic value.[8] Additionally, SLU recognise that the scientific research data generated by the university is of potential value to future SLU and other researchers. As such, they are aware that not only the data itself must be preserved but also its context and authenticity.

Regulatory Environment

SLU must comply with the Principle of Public Access to Official Documents. This legislation is similar to the Freedom to Information legislation in the UK.[9] As such, they must provide access to any public records – whether paper or digital – requested by the public in a timely manner.

Preservation Activity

Policies and Strategies

The SLU archives are active on a national level in Sweden and also on an international level through their involvement with the International Council of Archives.[10] The SLU archives find their involvement in national and international organisations extremely helpful for keeping up to date with developments in the field of digital archives and preservation. However, the SLU have only recently begun work in earnest on the development of an overall digital preservation strategy for the university. Participants state that previous efforts with regards to digital preservation have been piecemeal and have lacked the required commitment from senior management to ensure implementation. The SLU archives have incorporated digital preservation elements into the new IT policy currently being drafted for the university.

Generally, the SLU archives are responsible for appraising and preserving the administrative, educational and research records for the entire university. While some information management policies apply across the entire institution, others depend on departmental requirements. The IT department has had some involvement in the development of university-wide policies as well.

Selection

SLU archives do not differentiate between digital or paper records when it comes to their selection. Records of value are retained regardless of their format. The selection process is carried out cooperatively between the university archives and individual faculty departments. No paper or digital records can be disposed of without the approval of the university archives under archival legislation.

Preservation

SLU archives are aware of external standards, best practices and guidelines for digital preservation and keep up to date with initiatives in the European Union, the United States, and Australia. The SLU archives follow recommendations set out by the National Archives of Sweden in choosing standards for the preservation and delivery of

---

[8] The legal value is explained among other reasons by the Principle of Public Access to Official Documents mentioned below.
[9] In principle, all individuals – whether citizens of Sweden or of another country – have the right to read official documents held by public authorities. However, this right is subject to two restrictions: 1) the general public is entitled to read only those official documents that are classified as public documents; 2) some public documents are secret. http://www.sweden.gov.se/content/1/c6/01/62/87/5626168f.pdf.
[10] International Council of Archives. http://www.ica.org/.

their digital records. Currently, the SLU archives recommend that departments utilise file formats that are software and hardware independent. For numeric or "simple" databases, flat files are created with associated metadata and a Document Type Definition (DTD) in eXtensible Markup Language (XML) is generated. For text documents, plain text is recommended if formatting is not important. If retaining the original formatting is important, Portable Document Format (PDF) is used. Email to be preserved may be encoded in XML, but SLU still print them to paper for their long-term preservation. Official documents are also printed to paper for long-term storage. Digital images are captured and stored as Tagged Image File Format (TIFF) files or in Joint Pictures Expert Group (JPEG) format with lossless compression for the archival master. Currently, SLU have no strategy for the long-term preservation of web resources or multimedia, but this will be examined through the collaborative pilot project to establish a joint digital archive which is described below. SLU retains redundant copies of digital records to be preserved to serve as backup versions. However, SLU archives consider digital data stored off-line to be at high risk for degradation or inaccessibility.

SLU is participating in a new collaborative pilot project to establish a joint digital archive for scientific research with thirteen other Swedish universities. The fourteen universities involved will share the costs for the project. The project consists of two main groups – the national steering group and the institutional project group. The national steering group is made up of six representatives, four of which are archivists from Linköpings University, Lund University, Växjö University, and SLU. Two librarians from Gothenburg University/BIBSAM and Uppsala University complete the steering group. The SLU project group comprises nine scientists (representing all of the faculties), two ICT staff, two librarians and the university archivist as a project leader.

The project began in May 2004 and will run until May 2005. The project will use six different scientific databases and will test both functional and bit preservation according to the Open Archival Information System (OAIS) Reference Model and within the DSpace environment. During the first phase of the project (May to October 2004) two databases will be migrated and tested. The databases to be migrated are longitudinal experiments on soil from the Department on Soil Sciences and the Swedish National Survey of Lakes and Running Waters from the Department of Environmental Assessment. Partners will also try to agree on descriptive and administrative metadata for the scientific records. The major problem they have encountered is establishing an appropriate level of metadata that suits the various scientific disciplines and also meets the needs of future 'consumers'. The SLU archives feel that Dublin Core is not sufficient for the archival requirements of the joint digital archive. Therefore, they aim to develop their own model for the preservation of scientific databases following international metadata initiatives as a basis. The SLU archives have looked at CEDARS, NLA, NEDLIB and the Preservation Description Information according to the OAIS Reference Model. The SLU anticipate that part of the metadata generation process will be automated for the pilot project. This pilot project will also look to clearly identify roles and responsibilities with regards to digital preservation. The national and institutional project groups will meet from October 28-29, 2004 to evaluate the first phase of the project.

Access

Most of SLU's digital information is accessible online via the Internet. For non-public records, access is based on internal routines and authority systems. Redundant copies are stored offline.

Compliance Monitoring

While the SLU has an internal audit department,[11] participants revealed that internal compliance with policies, standards and other requirements are not currently audited on a regular basis. This is largely due to ambiguity over roles and responsibilities regarding the long-term preservation of SLU's digital data. This is something that will be investigated during the pilot project to establish a university-wide IT/digital preservation strategy for the SLU.

Digital Preservation Costs

SLU did not indicate what percentage of their budget is currently spent specifically on digital preservation. However, the SLU archives and IT department are adamant that the digital preservation budget must be increased in the future. It is hoped that the pilot project on digital preservation will illuminate the need for funding in this area. If and when more funding is available, SLU archives would like to see more spent on building an archival system and implementing the necessary organisational changes that will be required to ensure the long-term preservation of the university's digital records.

Future Outlook

SLU archives would like to see more work done on revising and adjusting internal preservation policies and strategies. They would also like to see additional resources going towards digital preservation and technical solutions. Their participation in the joint digital archive pilot project should help to ensure that both the SLU's digital resources and those of the wider academic community are accessible for the long-term.

---

[11] SLU Internal Auditing: http://www.slu.se/internrevision/indexeng.html.

Chapter 7: Conclusions

Digital preservation of scientific data is a problem for many universities and research organisations. SLU have lacked an institutional policy for the long-term preservation of their digital resources supported and coordinated by senior management. However, SLU have overcome this obstacle by looking outwith the university and collaborating with colleagues from other institutions. This proactive approach should help to ensure that SLU's digital resources are available in the long term. In fact, this joint pilot project could prove an integral turning point for the SLU in focusing the attention of senior management to the long-term care of the university's digital resources. This study illustrates the value of collaboration when facing the challenges of digital preservation, and the results of the pilot project will be of great interest not only for their technical findings but also as a model for potential economic sustainability. The results of this pilot project will be of great value to both the academic community in Sweden and the scientific research community in general.

CONTACT DETAILS

ERPANET Coordinator
George Service House
11 University Gardens,
University of Glasgow
Glasgow, G12 8QQ,
Scotland

Tel: +44 141 330 4568
Fax: +44 141 330 3788
Coordinator@erpanet.org

ERPANET STAFF

d i r e c t o r s
Seamus Ross, Principal Director
Niklaus Bütikofer, Co-Director
Mariella Guercio, Co-Director
Hans Hofman, Co-Director

c o o r d i n a t o r
Peter McKinney

e d i t o r s
Andreas Aschenbrenner
Georg Büchler
Joy Davidson
Prisca Giordani
Francesca Marini
Maureen Potter

www.erpanet.org

ELECTRONIC **R**ESOURCE **P**RESERVATION AND **A**CCESS **N**ETWORK