



Information Society
Technologies

erpastudies

ISSN
international
centre



erpastudies

www.erpanet.org

ERPANET – Electronic Resource Preservation and Access Network – is an activity funded by the European Commission under its IST programme (IST-2001-3.1.2). The Swiss Federal Government provides additional funding.

Further information on ERPANET and access to its other products is available at <http://www.erpanet.org>.

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server (<http://europa.eu.int>).

ISSN 1741-8682
© ERPANET 2004

Table of Contents

Table of Contents	1
Executive Summary.....	2
Chapter 1: The ERPANET Project.....	3
Chapter 2: Scope of the Case Studies.....	4
Chapter 3: Method of Working	6
Chapter 4: the ISSN International Centre	7
Chapter 5: Details and circumstances of the interviews.....	8
Chapter 6: Analysis	9
Perception and Awareness of Digital Preservation	9
Preservation Activity.....	9
Compliance Monitoring.....	11
Digital Preservation Costs	11
Future Outlook.....	11
Chapter 7: Conclusions	13
Appendix 1: References.....	14

Executive Summary

The ISSN International Centre is responsible for the International Standard Serial Number (ISSN), a key identifier in the world of serial publications. The ISSN textual catalogue stored in the ISSN Register database has been accumulating information since the 1970s, and it is destined to remain reference far into the future. Due to the Centre's circumspect maintenance of the system, its migration cycles can be kept rather sparse; the first major modernisation is currently in progress, after the system was operational reliably for thirty years. Along with the hardware environment also the software format of the ISSN records is currently being updated to the internationally acknowledged MARC 21 format. The ISSN network has always promoted standardisation and international cooperation. Currently, international consensus is forming on the unique and persistent identification of objects in the digital world, which also assists their long-term preservation. Being part of this effort, the ISSN number may become an even more crucial component in the world information infrastructure of the future.

Chapter 1: The ERPANET Project

The European Commission and Swiss Confederation funded ERPANET Project¹ (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.ermanet.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

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/www/content/documents/Digitalpreservationcharterv4_1.pdf.

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.

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 instrument³ 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 www.erpanet.org. 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: the ISSN International Centre

The International Standard Serial Number (ISSN) is an internationally recognized eight-digit standard identification number for serial publications. 'Serials' refers to periodicals, newspapers, proceedings, and any other publications in successive parts, issued on any medium including electronically.

Assignment of ISSN numbers is done by a continually expanding network of 75 national centres covering 239 countries. Coordination within the ISSN Network and policy making is assumed by the ISSN International Centre, which is located in Paris, France. Seventeen people work at the ISSN International Centre, yet in the whole network there are well over a hundred people involved.

The ISSN programme was developed under the framework of UNESCO's world science information system.⁵ In 1975 the release of the ISO International Standard 3297 defined the ISSN number.⁶ The ISSN International Centre is governed by representatives from UNESCO, the French government and from ISSN member states. They also provide the financial resources of the Centre. In addition, the Centre earns revenues from the sales of its publications and services, foremost the ISSN Register.

The ISSN Register is the compilation of all assigned ISSN numbers world-wide. This database is maintained by the International Centre. Formerly known as the International Serials Data System, the ISSN Register was established in 1974. Only now a completely new system is being installed, which is expected to launch in the end 2004.

⁵ World Science Information System - UNISIST, initiated in 1972 by the UNESCO General Conference. <http://mist.univ-paris1.fr/memoire/1998/afleury.html>.

UNESCO - United Nations Educational, Scientific and Cultural Organization;
<http://www.unesco.org/>.

⁶ ISO - International Organization for Standardization; <http://www.iso.ch/>.
ISO International Standard: Information and documentation - International Standard Serial Number (ISSN). Third Edition, 1998-11-01. ISO 3297:1998(E);
<http://www.niso.org/standards/resources/ISO3297-1998.pdf>.

Chapter 5: Details and circumstances of the interviews

Contributions came from the information management, senior management, and information technology areas at the ISSN International Centre in Paris.⁷ Interviews were conducted via phone and email in October 2003. In addition, publicly available documentation was consulted.

ISSN International Centre
20, rue Bachaumont
75002 Paris - France

⁷ We are thankful to Pierre Godefroy, Françoise Pellé, and Bojnourdi Kaikavous for their time.

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

The ISSN Register database is the heart of the organisation. The stewardship and preservation of its contents (the ISSN numbers and the corresponding data) constitute the organisation's core mission. Staff at the ISSN Centre are well aware of the value of the Register and their mandate.

Asset value and risk exposure

The Register is a dynamic database where information is not overwritten or even deleted. In 2003 it contained almost 1,100,000 entries, with an annual growth of up to 50,000 records and updates to some 100,000 records.

The most immediate preservation challenges lie in maintaining the active system, and in applying due diligence when migrating to new technology generations. While the latter is a specifically difficult task for a dynamic database, it is fortunately not required as often as one might expect. The last system has been active since the ISSN's inception in the 1970s, and it is updated only now after being active for 30 years.

Regulatory Environment

The ISSN Centre is an intergovernmental organisation and therefore not subject to any national laws. Even though the organisation voluntarily adheres to laws of the 'host' country France where applicable, there are no formal legal obligations to preserve the Register over the long term. However, as described above, this task is part of the organisation's mission.

Preservation Activity

Policies and Strategies

While ISSN staff are very much aware of the value of the ISSN Register, the centre lacks a formal long-term preservation policy at this stage. At the same time, however, the centre attaches great importance to the stewardship and to the safety of their active data. Also, the current migration of the ISSN system is being handled with great care: planning started in 2002 and the system is expected to go live by the end of 2004.

The ISSN network also collaborates with international stakeholders and adheres to international standards, which is conducive to interoperability and the stability of their information assets. The ISSN definition is an ISO standard itself, and the ISSN Register is actively working towards harmonising cataloguing rules. To this end they

are contributing to workgroups at IFLA, to the Anglo-American Cataloguing Rules⁸, and others. Results of these cooperations will be implemented in the new system.

Selection

The definition of a serial publication is defined in the ISO 3297 standard, the ISSN defining document. A serial publication as defined by this ISO standard refers to all publications in successive parts, and also includes serial publications in digital form. The respective national centre of the country where a serial is published is responsible for verifying the selection criteria. If applicable the national centre assigns an ISSN number and creates the ISSN record, which is transferred to the ISSN Register. The quality of the ISSN records is ensured through templates and format guidelines. Compliance to these is (rigorously) supervised by the ISSN Centre.

Preservation

The records in the ISSN Register database are accessible in MARC format, a national as well as international standard.⁹ MARC, short for MACHine Readable Cataloguing, is an encoding format used for storing bibliographic data. There are a number of MARC 'flavours', each with slightly different formats. ISSN originally created its own ISSN MARC, which is derived from USMARC, the original MARC format created at the Library of Congress.¹⁰ In the ongoing move to the new system the ISSN Register is now being transferred to MARC 21 format.¹¹ MARC 21 is more detailed and is subject to more frequent changes than ISSN MARC. The loss in stability, however, is balanced by the fact that MARC 21 has become the most prevalent MARC format. The transfer is therefore done for compliance and interoperability reasons.

The system currently active at the ISSN International Centre in Paris is for all intents and purposes still the first one installed. With some updates it has been working efficiently throughout the period. In the meantime, however, maintenance of the legacy system has become costly. In tandem with the need to install some fundamentally new features this triggered the decision to acquire a new system. This system is based on a commercial integrated library system, one of the six major products in this area on the international market.¹² The system is tailored to the specific needs of ISSN, and it will cater for Unicode¹³ and other current features.

While in the current system all data are stored at the ISSN centre in Paris, the new system will out-source storage. The out-sourcing partner is expected to ensure system reliability and data security. Negotiations of the contract terms are ongoing, but it appears most likely that the technology infrastructure will be provided by a public data service centre and the ISSN Centre will administer the system.

⁸ Anglo-American Cataloguing Rules - AACR; <http://www.nlc-bnc.ca/jsc/index.html>.

⁹ MARC - MACHine Readable Cataloguing; <http://lcweb.loc.gov/marc/>.

The MARC structure is constituted in international standards:

Format for Information Exchange (ISO 2709), as well as Information Interchange Format (ANSI Z39.2).

¹⁰ Library of Congress, United States of America: <http://www.loc.gov/>.

¹¹ Library of Congress – Network Development and MARC Standards Office: MARC 21 Concise Format for Authority Data. <http://www.loc.gov/marc/authority/>.

¹² The new system is a product by VTLS, called Virtua ILS – Integrated Library System; <http://www.vtls.com/>.

¹³ Unicode is a comprehensive character encoding standard defined by ISO/IEC 10646, a superset to the ASCII character set;

UNICODE - <http://www.unicode.org/>.

ASCII - <http://en2.wikipedia.org/wiki/ASCII>.

The database itself is not particularly large in terms of absolute storage space. This is due to the fact that the average record size is about 0,4 kilobyte. For the 1,100,000 entries contained in the database by 2003, this results in a total size of the ISSN Register of 450 megabytes. The average record size is expected to grow with the conversion from ISSN MARC to MARC 21, but it will still be a relatively small database.

Backups of the entire ISSN Register are being secured on magnetic tape on a daily basis. Tapes are stored at a secure place, and they are not reused in order to avoid media deterioration.

Access

Access to the ISSN Register is possible in two ways. Web access is provided via ISSN Online on a subscription basis,¹⁴ and a CD is distributed to more than 250 subscribers worldwide on a quarterly basis. The production of the CDs is outsourced to a small American company.

Each CD contains the entire ISSN Register in a MARC format. These CDs can be regarded as a state history of the ISSN Register. Accessing CDs with earlier versions of the ISSN Register is expected to be possible due to the simple format of the ISSN records. There is, however, no policy for preserving the old CDs themselves.

Compliance Monitoring

As there is no formal preservation policy in place at this point, compliance benchmarks have not yet been established. The ISSN Centre, however, cares for premier quality assurance. Automatic backups, for example, are monitored and the corresponding log files are inspected. The quality of the records conversion from ISSN MARC to MARC 21 will be ensured along similar lines.

Digital Preservation Costs

Actions that support the long-term preservation of the ISSN Register cannot be separated from other operational costs, including system maintenance. The ISSN Centre attaches great importance to the adequacy of its systems, and dedicates a substantial financial effort to the acquisition as well as the annual maintenance of its new system.

Future Outlook

In their effort to provide a timely service, the ISSN Centre may consider providing their MARC 21 records in MARC XML,¹⁵ a new presentation of the MARC format that has recently gained popularity. At the moment, however, the move to the new system demands all their resources.

The ISSN is of particular interest to digital preservation, as it might provide a core building block in the persistent identification of digital serials. The ISSN Centre made the suggestion¹⁶ of using the ISSN as a URN (Uniform Resource Name)¹⁷ identifier. Any electronic serial with an ISSN number assigned could then be uniquely identified

¹⁴ ISSN Online; <http://online.issn.org/>.

¹⁵ Library of Congress: MARC in XML. <http://www.loc.gov/marc/marcxml.html>.

¹⁶ IETF RFC 3044 (see Appendix 1: References)

¹⁷ K.Sollins, L.Masinter (Network Working Group): IETF RFC 1737 (Internet Engineering Task Force - Request for Comment) - Functional Requirements for Uniform Resource Names. December 1994. <http://www.ietf.org/rfc/rfc1737.txt>.

and accessed via an ISSN URN resolution system. An ISSN URN Demonstrator has been implemented proving the viability of such an approach.¹⁸

Lastly, the ISSN Network is open to take any steps that may enhance their aptitude to preserve the ISSN Register into the far future. To this end the network upholds various collaborations and seeks ever new partnerships.

¹⁸ ISSN URN Demonstrator; <http://urn.issn.org/>.

Chapter 7: Conclusions

Discussions with the ISSN team highlighted the difficulty of preserving dynamic and open databases, confirming the conclusions from the ERPANET workshop in Bern, April 2003.¹⁹ As the ISSN Register approaches their task, the preservation of an active database consists of two processes: phases of maintaining the active systems and a sort of rolling migration, intermitted by more fundamental changes to the system infrastructure. By diligently maintaining the active systems and including long-term thinking at system design stage, the ISSN Centre successfully avoided frequent fundamental updates of their systems, and thereby lowered risks and saved overall costs.

The fact that the ISSN Centre is conducting their first fundamental update only now after 30 years of adequate operations is through the textual format of the content. This combination of basic systems and data formats, as well as diligent maintenance and long-term thinking – this pragmatic stance may be a formula of success for many other organisations as well.

Another important component to this is the stability of the format of the ISSN records. The ISSN Centre was able to maintain ISSN MARC with relatively sparse updates. The organisation hopes to maintain this format stability with the future format employed, the international standard MARC 21. In addition to the stability of standard formats, interoperability, benefits from the communal work, and market pressure are other strong arguments for complying with them.

¹⁹ See the report from the event:
http://www.erpanet.org/www/products/bern/Bern_Report_final.pdf.

Appendix 1: References

The website of the *ISSN*,
<http://www.issn.org>

ISO International Standard 3297: Information and documentation - International Standard Serial Number (ISSN).
<http://www.iso.ch/>.

Pierre Godefroy: Call for tender for the ISSN data system. 2002.
<http://www.issn.org:8080/English/pub/consultation/>.

Françoise Pellé: ISSN International Centre – Progress Report. Conference of Directors of National Libraries (CDNL); 21 August 2002.
<https://www.nla.gov.au/initiatives/meetings/cdnl/issn.html>.

S. Rozenfeld (ISSN International Centre): Using The ISSN (International Serial Standard Number) as URN (Uniform Resource Names) within an ISSN-URN Namespace. IETF RFC 3044 (Internet Engineering Task Force - Request for Comment); January 2001.
<http://www.ietf.org/internet-drafts/draft-rozenfeld-urn-issn-00.txt>.

MARC - MACHine Readable Cataloguing; <http://lcweb.loc.gov/marc/>.

Library of Congress – Network Development and MARC Standards Office: MARC 21 Concise Format for Authority Data. <http://www.loc.gov/marc/authority/>.

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

directors

Seamus Ross, Principal Director
Niklaus Bütikofer, Co-Director
Mariella Guercio, Co-Director
Hans Hofman, Co-Director

coordinator

Peter McKinney

editors

Andreas Aschenbrenner
Georg Büchler
Joy Davidson
Prisca Giordani
Francesca Marini
Maureen Potter

www.erpanet.org