

ELECTRONIC RESOURCE PRESERVATION AND ACCESS NETWORK



ERPANET@SOA2004: Introducing ERPANET's Knowledge Base

Seamus Ross, Hans Hofman, Peter McKinney

ERPANET

Simply:

ERPANET 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.

ERPANET EU FP5 Project

ERPANET aims to do the following:

- **raise awareness of preservation** needs and strategies;
- **appraise and evaluate information sources** and developments and increase the accessibility of this work;
- provide an **EU-wide enquiry and advisory service**;
- organise 15 developmental **seminars and workshops** **stimulate research** in digital preservation;
- **promote awareness** among software producers of the preservation needs of the user community; and
- research to **improve method and practice**.

Who is ERPANET?

Institutions:

- The Humanities Advanced Technology and Information Institute (HATII)
- Schweizerisches Bundesarchiv
- Nationaal Archief van Nederland
- Institute for Archival and Library Science (Università degli Studi di Urbino)

The Management Team

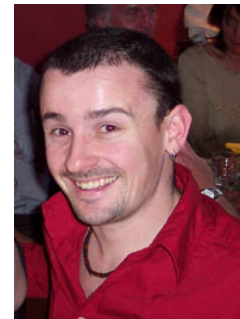
The Directors

- Niklaus Bütikofer (Schweizerisches Bundesarchiv)
- Mariella Guercio (ISTBAL, Università di Urbino)
- Hans Hofman (Nationaal Archief Nederland)
- Seamus Ross (HATII, University of Glasgow)



The Co-ordinator

- Peter McKinney (HATII, University of Glasgow)



The ERPANET Editors

HATII, University of Glasgow

– **Joy Davidson & Adam Rusbridge**

ISTBAL, Università di Urbino

– **Prisca Giordani & Francesca Marini**

Nationaal Archief Nederland

– **Andreas Aschenbrenner & Maureen Potter**

Schweizerisches Bundesarchiv

– **Georg Büchler & Niklaus Bütikofer**

Organisations Need Help

- Off-the-shelf policy statements
- Business cases & strategies
- Digestible guidance on technologies and their preservation implications
- Improved models (reference, costs, standards, functional requirements)
- Simple Guidelines on digital survival
- Guidance on creating data repositories
- IPR support and guidance



Maria Guercio and John McDonald at the *Preservation Policy and Procedures* co-sponsored with the Archives de France – Le Centre des Archives Contemporaines, Fontainebleau, January 2003.

Our Response

erpaPRODUCTS

erpaEvents

Workshops and Seminars

erpaGuidance

Selecting technologies, Risk management, Costing, Policies

erpaStudies

Investigation of practice and solutions

erpaCharter

Principles of digital preservation

erpaAdvisory

Free Advisory service

erpaAssessments

Analytical commentaries

erpaDirectory

List of international projects and resources

erpaEprints

Open Archives repository

erpaDiscussion

erpaChat and Forums

erpaEvents Thematic Workshops & Seminars

- Exploring core topics with experts and specialised institutions
- Increasing the level of awareness within the communities of digital information providers, practitioners, experts and preservers
- Developing basic recommendations
- Providing documentation of new research directions and results
- Provide skills development opportunities to the community



Speakers at the Lisbon Seminar on *The Selection, Appraisal and Retention of Digital Scientific Data*,

Workshops & Seminars

- Hosted 20 Seminars/Workshops over 36 months
- Workshops & Seminars

-XML for digital preservation (Urbino, October 2002)

-Introduction to the OAIS Model (Copenhagen, November 2002)

-Preservation Policy and Procedures (Fontainebleau, January 2003)

-Metadata in Digital Preservation, (Marburg, September 2003)

-Digital repositories, (Rome, November 2003)

-Preservation of Scientific Data Sets (Lisbon, December 2003)

-Audit and Certification (Antwerp, April 2004)

-File Formats (Vienna, May 2004)

-Persistent Identifiers (Cork, June 2004)

-Business Models (Amsterdam, September 2004)

-Preserving Digital Art (Glasgow, October 2004)

-Workflow (Budapest, October 2004)

-Managing and Archiving Records in the Digital Era, (Bern, October 2004)

Over 35 sponsoring organisations

- **Ministero per l'università e la ricerca (Italy)**
- **Cambridge University and Library**
- **MIT (USA)**
- **CODATA**
- **Trivadis AG**
- **City Archives of Antwerp**
- **National Library of Austria**
- **University of Cork**
 - **Archives de France – Le Centre des Archives Contemporaines (Fontainebleau)**
 - **The Biblioteca Nacional España**
 - **Caja Castilla la Mancha**
 - **Det Kongelige Bibliotek (Denmark)**
 - **Archivschule Marburg**
 - **Accademia dei Lincei**



Participants at the *File Formats* Seminar in Vienna, May 2004

BENEFITS of Events

- Creation of authoritative sources of information
- Fora for information and preservation professionals that can be extended into their own and other communities
- Exchange of experiences; successes and failures
- Post-event analytic report recommending areas for research, development and service provision following each event – substantial interest in these reports (e.g. nearly 12000 downloads of Lisboa Report since April 2004)
- erpaChat sessions



Mr Allemann, Trivadis AG
Switzerland co-sponsor of the
workshop on *Long term
preservation of databases*, April
2003.

Some Lessons learned

- Community hungry for practical case studies and reports of real world experiences
- Lack of good examples makes it difficult to identify speakers with expertise, technical knowledge of a high enough calibre and suitable presentation skills
- Audience has a high and increasing skill-base
- Discussion with experts through break-out sessions in demand
- Cross-domain audience helps to bring down barriers between domains
- Evaluations show that ERPANET events are in demand and participants in one event intend and do return to others



ERPANET Bern Workshop on *Long term preservation of databases*, April 2003.

erpaAssessments

Authoritative online information services on Digital Preservation

- Over 250 commentaries
- Analytic Reviews of multi-lingual published and online literature
 - including assessment of 100 journals
- Assessment of technical guidelines and relevant standards
- Projects reviewed

erpaGuidance

- Policy Building Framework
 - (2070 downloads Aug03-Jul04)
- Costing Digital Preservation Strategies
 - (2287 downloads Aug03-Jul04)
- Digital Object Risk Management Methods
 - (1805 downloads Aug03-Jul04)
- Selecting Technologies
 - (1478 downloads Aug03-Jul04)
- Ingest Strategies

erpaEprints

- Ensuring accessibility to digital preservation literature, both grey and published
- Creating a single point of access
- Better profiling of the literature
- Researching usage of online repositories



Andrew Wilson of the National Archives of Australia speaking *Metadata for Digital Preservation*, (September 2003).

erpaAdvisory

- Provides the community with access to answers to digital preservation expertise
- Building a resource of authoritative answers to digital preservation questions
- Making accessible examples of best practice
- Currently under-utilised



Robert Kahn (Corporation for National Research Initiatives) and Tito Orlandi (Centro Linceo Interdisciplinare) at *Trusted Repositories for Preserving Cultural Heritage* (Rome, November 2003)

erpaStudies their aims



Stuart Weibel of OCLC at the grave of George Boole following the ERPANET seminar on *Persistent Identifiers*, Cork, June 2004

- Build a picture of current methods matched against their context of use;
- Underpin a definition of best practice;
- Accumulate and make accessible information about approaches to digital longevity;
- Produce materials for seminars and workshops;
- Enable comparisons of approaches across sectors;
- Enable the development of guidance documents; and,
- Identify issues for further research.

erpaStudies parameters

- Case studies in public and private sectors
- Aim to achieve spread across organisational size and location
- Diversity of organisational type, activity, regulatory framework, and culture
- Various typed of records / objects
- Outcomes:
 - Vertical View: reviews of how different sectors approach digital preservation
 - Horizontal View: looking across a range of topics what are best practices across domains.

erpaStudies methodology

- Research by questionnaire and interview (normally by telephone)
- Examined how other projects had conducted surveys to inform our own design
 - The Pittsburgh Project and InterPARES
- Designed focused template which we hope other projects will adopt when conducting similar studies
- Interviews with three staff members sought: Business Manager, IT Manager, Archivist/Records Manager

The Sample

- Almost 500 organisations contacted
- Current participation rate: just about 15% (65-70); final result will be about 100
- Many initially said ‘yes’ or ‘maybe’ and then ‘no’
 - “currently developing a digital preservation programme and at an early stage”
 - “no time and capacity for participation”
 - “too early to participate”
 - Pulled out when seeing questionnaire: common excuse, no one able to complete form
 - Dropped out after reading report (e.g. BNF, International Court of Justice) - occasionally on legal advice

The Participating Organisations

- 1st Round Participants (20 organisations)
 - Astra Zeneca PLC, Aventis Pharma Germany GmbH, Bayer AG, Böhringer Ingelheim, GlaxoSmithKline plc (GSK), Organon NV, Pfizer, RAI (Radiotelevisione Italiana), BBC (British Broadcasting Corporation), ERT (Elliniki Radiophonia Tileorassi SA), SF DRS (Schweizer Fernsehen der Deutschen und Rätoromanischen Schweiz), Gruppo Mondadori (), Hachette Filipacchi Médias, HMSO (Her Majesty's Stationary Office), Kluwer Academic Publishing, Office for Official Publications of the European Communities (Publications Office), Oxford University Press, Random House, Swisscom, Orange
- 2nd Round Participants (35 organisations)
 - Migros, Co-OP, Loeb (all Switzerland), Kesko (Finland), AB Vasilopoulos (Greece), Canadian Institute for Historical Microreproductions, CINECA (Italy), Edinburgh University Library (UK), SDA, European Patent Office (Netherlands), Banking House, Bio-inspecta Council of Europe (Luxemburg), International Labour Organisation, Universal Postal Union (Switzerland), WIPO (Switzerland), ISS, MAIB (UK), NHDA (Netherlands), Office for Metropolitan Architecture (Netherlands), TESSELLA (UK), Steel Component company, Meteorological Institute (UK), Meteorological Institute (Netherlands), TATE, National Archives Scotland (UK), Nederlandse Aardolie Maatschappij (Netherlands), Dutch Kadaster (Netherlands), Centraal Bureau voor de Statistiek (Netherlands), Theater Instituut Nederland (Netherlands), Ferrovie dello stato (Italy), Bank of Italy, British Telecom (UK), Swedish Agricultural University (Sweden), National Library of Wales (UK).

Focus of Investigations

- Perception and awareness of risk posed by digital materials
- Understanding how digital preservation affects the organisation
- Identifying what actions have been taken to prevent loss of digital materials
- Develop an understanding of the way preservation activities are monitored (policies or programmes)
- Understand how organisations plan future preservation requirements (e.g. costs)
- What technologies are in place

General Findings (1)

- Variation in awareness of risk
- Value of information depends upon recognition of business/organisational dependency, potential of re-use, or risk associated with information
- Responsibility rarely taken at corporate level
- Diversity of strategies, but few have them
- If not explicitly defined in policy, it will be unlikely digital preservation is really achieved
- Activity is fragmentary: practices tend to be *incomplete, ad hoc, and individual*

General Findings (2)

- Recognised need for better sharing of knowledge
- Organisations are often waiting for external solutions
- Collaborative sectoral action rare (broadcasting is an exception), but if, then increased confidence and action
- Selection criteria from the paper world have been transferred to the digital environment
- Preservation and storage poorly understood
- Costs difficult to justify in the corporate environment

General Findings (3)

- Lack of a single point of contact—no one knows who is responsible, or has the knowledge
- Many approaches are IT-driven, because of lack of ‘archiving’/preservation advice,
 - Due to lack of influence or knowledge of records manager?
- Profile of records or archives department needs to be raised

Objectives Achieved

- Accumulated materials to illustrate methodologies and practices
- Created a basis for establishing practical guidelines
- Raised awareness among almost 500 institutions and individuals
- Stimulated sharing
- Added objective understanding to a profile which has often been based on single analyses and anecdotal

The Follow-up (1): What are we doing with the results?

- Sharing knowledge: website, articles and conferences
- Attempting to use our knowledge to address some of their problems
 - e.g. developing tools
 - Business cases and strategies
 - Digestible guidance on technologies and their preservation implications
 - Construction of improved models (reference, cost...)
 - Guidance on repository design, maintenance, audit

The Follow-up (2): What are we doing with the results?

- Conducting Further Research, for example
 - ERPANET's Portable Object Ingest Project (POIP)
 - Digital Curation Centre's research on ingest and extraction of semantic metadata (UK)
- Identify topics and organise seminars and workshops

The Website – Relunched 24/8/04

- Establish an extensive Web-based service
- Maintain and expand website and access to products
- Identify and agree evaluation and QA criteria for appraisal of information
- Design and pilot templates for data collection
- Completely redesign the site—release 27th Aug 2004

Project Management

- Distributed projects exceptionally hard to manage
- Meetings
- Weekly conference calls
- Focus on establishing processes and procedures
- Quality control

What Does the Future Hold?

- Uncertain
 - Seeking additional funding streams
 - Investigating self-funding options
- We believe though that
 - ERPANET has created a strongly collaborative network
 - ERPANET has established a mechanism to deliver high quality resources, services, and continuing professional development opportunities
 - ERPANET offers an international model of capacity building and information sharing

VISIT The ERPANET WEBSITE:

<http://www.erpanet.org>