

The Canadian Conservation Institute (CCI) considers the following information to be useful and relevant for conservation research or reference purposes. This content has been provided here as archived material, which means it is not subject to Government of Canada Web Standards. To request an alternate format, please contact CCI ([www.cci-icc.gc.ca](http://www.cci-icc.gc.ca)).

L'Institut canadien de conservation (ICC) considère que les renseignements suivants sont à la fois utiles et pertinents pour la recherche en conservation ou à des fins de référence. Ce contenu a été fourni ici à titre de matériel archivé, ce qui signifie qu'il n'est pas assujéti aux normes Web du gouvernement du Canada. Pour obtenir une version dans un autre format, veuillez communiquer avec l'ICC ([www.cci-icc.gc.ca](http://www.cci-icc.gc.ca)).

CANADIAN  
CONSERVATION  
INSTITUTE



INSTITUT  
CANADIEN DE  
CONSERVATION

## Preservation of Electronic Records: New Knowledge and Decision-making

## La préservation des documents électroniques : Information récente et prise de décisions



## Table of Contents/Table des matières

Foreword / Avant-propos .....	i
Organizing Committee / Comité organisateur .....	ii
Acknowledgements / Remerciements .....	iii
Preface / Préface .....	v
Introduction: <i>Choices and Decisions</i> .....	1
Tom Strang	
<b>Appraisal of Electronic Records/Évaluation des documents électroniques</b>	
<i>Appraising Digital Materials for Preservation as Cultural Heritage</i> .....	7
Terry Eastwood	
<i>Preserving Digital Documentary Heritage in Libraries: What Do We Select?</i> .....	13
Ross Harvey	
<i>Selecting for Survival: Developing a Model for Selecting Audiovisual Archival Documents for Preservation Reformatting</i> .....	21
Nancy Marrelli	
<b>Authenticity of Electronic Records/Authenticité des documents électroniques</b>	
<i>Authenticity Requirements for Electronic Records</i> .....	25
Luciana Duranti	
<i>Authenticity and Accessibility, Objects and Technologies: Digital Preservation Basics</i> .....	31
Bruce Walton	
<b>Developing a Preservation Strategy for Electronic Records/ Élaboration d'une stratégie de préservation des documents électroniques</b>	
<i>Is Educating Archivists Enough?: Assessing the Impact of the ACA Institute 2002, "Approaches to the Preservation of Electronic Records"</i> .....	39
Amy Marshall and Wendy Duff	
<i>Preserving Digital Objects in Small- to Medium-sized Institutions</i> .....	47
Bill Walker	
<i>Developing a Digital Preservation Strategy at The British Library: Application of the Handbook Preservation Management of Digital Materials</i> .....	55
Deborah Woodyard and Helen Shenton	
<i>Digital Preservation: Best Practice for Museums</i> .....	61
Tim Au Yeung	
<i>Net Art et conservation, quelles stratégies pour le musée?</i> .....	71
Anne Laforet	
<i>Record, Play, Fast Forward: Developing Strategies for the Care of Electronic Media Art at the Art Gallery of Western Australia</i> .....	77
Vanessa Griffiths	
<i>Preserving Research Data: A Time for Action</i> .....	83
Charles Humphrey	
<i>Archival Science and Archival Engineering: Building a New Future for the Past</i> .....	93
Kenneth Thibodeau	
<i>Digital Preservation Strategies: The Initial Outcomes of the ERPANET Case Studies</i> .....	99
Seamus Ross, Monica Greenan, and Peter McKinney	

<b>Preservation Strategies for Electronic Records/Stratégie de préservation des documents électroniques</b>	
<i>Implementation of the Victorian Electronic Records Strategy (VERS) within the Government of the State of Victoria, Australia</i> .....	115
Howard Quenault	
<i>Volatile Memory</i> .....	121
Ulrich Lang	
<i>Robert Creeley's Computer Files: A Collaborative Approach to link Access with Preservation</i> .....	127
Michael G. Olson	
<i>Old Wine in New Skins: Migrating Legacy Data in a Museum Context</i> .....	135
Richard Gerrard and David Harkness	
<i>A Strategy for Archiving Web Sites at Library and Archives Canada</i> .....	143
Susan Haigh and Roselyn Lilleniit	
<i>Preserving Audiovisual Records of the University of Manitoba Archives and Special Collections</i> .....	149
Jane Dalley	
<i>Preservation Strategies for Digital Collections: A Critical Review</i> .....	155
Franziska S. Frey and Stephen Chapman	
<i>Computer-based Sound Recording Preservation and its Application in Small Communities</i> .....	157
Norm Glowach and John Poirier	
<i>Preservation and the Principles of Access</i> .....	165
Alexander Eykelhof	
<i>Archiving Electronic Media</i> .....	171
Jim Wheeler	
<b>Media Knowledge/Connaissance des supports</b>	
<i>Preservation of Electronic Records: Status of ISO Standards</i> .....	181
Peter Z. Adelstein	
<i>Waiting to Happen: Lessons from Preserving Disaster-afflicted Electronic Media in an Archival Collection</i> .....	189
Hannah Frost	
<i>Longevity of CD Media: Research at the Library of Congress</i> .....	197
Chandru J. Shahani, Basil Manns, and Michele Youket	
<i>Restless Records: The State of Small Video Art Archives in Canada</i> .....	207
Maija Martin	
<b>Posters/Affiches</b>	
<i>Survey Results: Preservation Management of Born-digital Documents in Manuscript Repositories in the United States</i> .....	217
Ruth Elizabeth Bryan	
<i>Virtual Remote Control: Preservation Risk Management for Web Resources</i> .....	219
Nancy Y. McGovern, Anne R. Kenney, Richard Entlich, William Kehoe, and Peter Botticelli	
<i>Role of Lacquer and Ink in Protecting Compact Discs exposed to Gaseous Pollutants</i> .....	221
Léon-Bavi Vilmont	

## Foreword

The preservation of electronic records is one of the latest challenges facing the conservation and heritage communities. It is a complex and multifaceted task that includes content, media, hardware, and software.

The program for *Symposium 2003 - Preservation of Electronic Records: New Knowledge and Decision-making* was developed to deal systematically and logically with the various issues. The organizing partners — the Canadian Conservation Institute (CCI), Library and Archives Canada (LAC), and the Canadian Heritage Information Network (CHIN) — all have extensive experience in the field of electronic record preservation, and each brought a unique perspective on specific aspects of the topic.

But the challenge of preserving electronic records extends well beyond the traditional heritage community. It really includes all custodians of electronic information — from corporations to government agencies to individuals. To meet this need, the symposium also included a separate half-day event for the general public.

The organizing partners were delighted to welcome more than 350 delegates to the symposium; of these, 85% were from Canada, 10% from the United States, and the rest from a variety of countries including Australia, Bermuda, the Cayman Islands, Cuba, France, Italy, Malaysia, Mexico, the Netherlands, New Zealand, Taiwan, and the United Kingdom. Everyone participated actively in the discussions, and returned to their institutions with not only a better understanding of the challenges but also with viable and practical solutions that can be implemented immediately.

*Charles Costain*  
Acting Director General  
Canadian Conservation Institute (CCI)

*Ian E. Wilson*  
National Archivist  
Library and Archives Canada (LAC)

*Roch Carrier*  
National Librarian  
Library and Archives Canada (LAC)

*Jean-Marc Blais*  
Acting Director General  
Canadian Heritage Information Network (CHIN)

## Avant-propos

La préservation des documents électroniques compte parmi les plus récents défis que doivent relever les collectivités de la conservation et du patrimoine. Cette tâche complexe concerne de nombreux éléments tels que le contenu, les médias, le matériel et les logiciels.

L'objectif du programme du *Symposium 2003 – La préservation des documents électroniques : Information récente et prise de décisions* est de se pencher systématiquement et logiquement sur les divers enjeux. Les organisateurs partenaires – l'Institut canadien de conservation (ICC), Bibliothèque et Archives Canada (BAC) et le Réseau canadien d'information sur le patrimoine (RCIP) – possèdent tous une vaste expérience de la préservation des documents électroniques, et chacun d'entre eux apporte un éclairage unique à des aspects précis de la question.

Mais le défi que représente la préservation des documents électroniques dépasse de loin la collectivité du patrimoine. Il inclut en fait tous les détenteurs de renseignements électroniques – des sociétés aux organismes gouvernementaux, et jusqu'aux particuliers. En vue de satisfaire ce besoin, une activité d'une demi-journée à l'intention du grand public fut présentée dans le cadre du symposium.

Les organisateurs partenaires ont été ravis d'accueillir plus de 350 participants, dont 85 % venaient du Canada, 10 % des États-Unis et 5 % de divers pays y compris l'Australie, les Bermudes, les îles Caïmans, Cuba, la France, l'Italie, la Malaisie, le Mexique, les Pays-Bas, la Nouvelle-Zélande, Taïwan et le Royaume-Uni. Tous ont pris part activement aux discussions, et tous sont retournés dans leurs établissements avec non seulement une meilleure compréhension des défis à relever, mais également avec des solutions pratiques qu'ils seront en mesure de mettre en œuvre sur-le-champ.

*Charles Costain*  
Directeur général par intérim  
Institut canadien de conservation (ICC)

*Ian E. Wilson*  
Archiviste national du Canada  
Bibliothèque et Archives Canada (BAC)

*Roch Carrier*  
Administrateur général / Bibliothèque nationale du Canada  
Bibliothèque et Archives Canada (BAC)

*Jean-Marc Blais*  
Directeur général par intérim  
Réseau canadien d'information sur le patrimoine (RCIP)

## Organizing Committee

**Co-Chairs:** Jane Down (CCI), Mary Murphy (LAC), Richard Green (LAC), Pat Young (CHIN)

**Program:** Joe Iraci (CCI) (*Chair*), Tom Strang (CCI), René Paquet (LAC), Bruce Walton (LAC), Gilles St-Laurent (LAC), Fay Hjartarson (LAC), Sheila Carey (CHIN)

**Posters:** Wanda McWilliams (LAC) (*Chair*), Alison Bullock (LAC), Janet Mason (CCI), Greg Hill (LAC), Bonnie Szirtes (CHIN)

**Publications/Signs:** Barbara Patterson (CCI) (*Chair*), Linda Leclerc (CCI), Sophie Georgiev (CCI)

**Tours:** Mary-Lou Simac (CCI) (*Chair*), Brenda Muir (LAC), Mike Thompson (LAC), Denis Gagnon (LAC)

**Preservation Quest:** Linda Street (CCI) (*Chair*), Joe Iraci (CCI), Lise Perron-Croteau (CCI), Carol MacIvor (CCI), Shanna Ramsay (CCI), Yvette Hackett (LAC), Louisa Coates (LAC), Randall Ware (LAC), Chris Dicks (LAC), Paul Lima (CHIN), Jane Down (CCI), Greg Hill (LAC)

**Local Arrangements:** Greg Hill (LAC) (*Chair*), Katherine Phillips (LAC), Natalie Labonté (LAC), Judith Falardeau (LAC), Colette Landry (CCI)

**Trade Show/Marketing:** Shanna Ramsay (CCI) (*Chair*), Chris Dicks (LAC)

**Registration/Correspondence:** Christine Bradley (CCI) (*Chair*)

**Finance:** Colette Landry (CCI) (*Chair*)

**Media Relations:** Carol MacIvor (CCI) and Louisa Coates (LAC) (*Co-Chairs*), Pauline Portelance (LAC), Paul Lima (CHIN), Shanna Ramsay (CCI)

## Comité organisateur

**Coprésidents :** Jane Down (ICC), Mary Murphy (BAC), Richard Green (BAC), Pat Young (RCIP)

**Programme :** Joe Iraci (ICC) (*président*), Tom Strang (ICC), René Paquet (BAC), Bruce Walton (BAC), Gilles St-Laurent (BAC), Fay Hjartarson (BAC), Sheila Carey (RCIP)

**Affiches :** Wanda McWilliams (BAC) (*présidente*), Alison Bullock (BAC), Janet Mason (ICC), Greg Hill (BAC), Bonnie Szirtes (RCIP)

**Publications/Panneaux :** Barbara Patterson (ICC) (*présidente*), Linda Leclerc (ICC), Sophie Georgiev (ICC)

**Visites :** Mary-Lou Simac (ICC) (*présidente*), Brenda Muir (BAC), Mike Thompson (BAC), Denis Gagnon (BAC)

**À la sauvegarde de votre patrimoine :** Linda Street (ICC) (*présidente*), Joe Iraci (ICC), Lise Perron-Croteau (ICC), Carol MacIvor (ICC), Shanna Ramsay (ICC), Yvette Hackett (BAC), Louisa Coates (BAC), Randall Ware (BAC), Chris Dicks (BAC), Paul Lima (RCIP), Jane Down (ICC), Greg Hill (BAC)

**Organisation locale :** Greg Hill (BAC) (*président*), Katherine Phillips (BAC), Natalie Labonté (BAC), Judith Falardeau (BAC), Colette Landry (ICC)

**Salon commercial/Promotion :** Shanna Ramsay (ICC) (*présidente*), Chris Dicks (BAC)

**Inscription/Correspondance :** Christine Bradley (ICC) (*présidente*)

**Finances :** Colette Landry (ICC) (*présidente*)

**Relations avec les médias :** Carol MacIvor (ICC) et Louisa Coates (BAC) (*coprésidentes*), Pauline Portelance (BAC), Paul Lima (RCIP), Shanna Ramsay (ICC)

## Acknowledgements

The hosting partners would like to extend their thanks to the Organizing Committee and all volunteers, speakers, poster presenters, session chairs, exhibitors, and sponsors, including the following organizations:

Audio Conservation Studio of Library  
and Archives Canada  
Audio Visual Preservation Trust  
Gatineau Preservation Centre  
Canada Museum of Science and Technology  
Canadian Conservation Institute  
Canadian Heritage Information Network  
Canadian Museum of Civilization  
Department of Canadian Heritage  
Library and Archives Canada  
The Willow Group  
Translation Bureau, Public Works and  
Government Services Canada  
Tunstall & Tunstall Data Recovery Services

## Remerciements

Les partenaires hôtes désirent remercier le Comité organisateur et tous les bénévoles, les conférenciers, les personnes qui ont présenté des affiches, les présidents de séances, les exposants et les commanditaires, y compris les organismes suivants :

Bibliothèque et Archives Canada  
L'Institut canadien de conservation  
Le Bureau de la traduction, Travaux publics  
et Services gouvernementaux Canada  
Le Centre de préservation de Gatineau  
Le ministère du Patrimoine canadien  
Le Musée canadien des civilisations  
Le Musée des sciences et de la technologie du Canada  
Le Réseau canadien d'information sur le patrimoine  
Le studio des enregistrements sonores de la Division  
de la musique (Bibliothèque et Archives Canada)  
Le Trust pour la préservation de l'AV  
Tunstall & Tunstall Data Recovery Services  
The Willow Group

## Preface

*Symposium 2003 - Preservation of Electronic Records: New Knowledge and Decision-making* took place in Ottawa on September 15–18, 2003, with a program that took the form of a decision tree. Following this format, the sessions could focus on areas that related to each other in a structured manner and all the key issues relating to the preservation of electronic records could be introduced in a logical sequence from the opening to closing speakers. The topics discussed included not only storage media but also the issues that must be addressed before considering how long an electronic record will last. [For the purpose of this symposium, the term “electronic record” was used to describe a large variety of records, e.g. audio and video recordings in analog or digital form and data files such as text, spreadsheets, e-mails, etc.]

The decision tree was initially formulated by the Program Committee, but refined by Tom Strang from CCI and Bruce Walton from LAC. Addressing the key decisions/choices that an institution needs to make when considering the acquisition and preservation of electronic records led to five main session categories: Appraisal of Electronic Records; Authenticity of Electronic Records; Developing a Preservation Strategy for Electronic Records; Preservation Strategies for Electronic Records; and Media Knowledge. The call for papers clearly outlined this approach, and resulted in the submission of more than 45 abstracts. Because of the decision tree program structure, the process to select the papers was very specific. The final program consisted of 29 papers from 6 countries: Canada (14); United States (8); Australia (3); United Kingdom (2); France (1); and Germany (1). In addition to the main program, 4 posters that fit into the program structure were also presented (1 from Canada, 2 from the United States, and 1 from France).<sup>1</sup>

Another objective of the program was to feature a wide variety of small- to medium-sized institutions that included not only archives and libraries, but also cultural institutions such as art galleries and museums that are faced with preserving electronic records. For example, art galleries often include video art in their collections, but the needs of these electronic records are quite different than the needs

---

1. One of the posters that was presented at the symposium was the decision tree that appears on p. 2 of the “Introduction” of this book of postprints. Hence, only three abstracts are included in the “Posters” section.

## Préface

*Symposium 2003 – La préservation des documents électroniques : Information récente et prise de décisions* a eu lieu à Ottawa, du 15 au 18 septembre 2003, avec un programme qui a pris la forme d’un arbre de décision. Selon ce format, les séances pouvaient porter sur des sujets liés entre eux de façon structurée et tous les points importants concernant la préservation des documents électroniques pouvaient être présentés logiquement, du premier au dernier conférencier. Les sujets discutés incluaient non seulement les supports de préservation, mais aussi les questions à traiter avant de se demander quelle est la durée de vie d’un document électronique. [Nota : Pour les fins de ce symposium, le terme « document électronique » a été utilisé pour désigner un large éventail de documents, comme les enregistrements audio et vidéo sous forme analogique ou numérique et les fichiers de données tels que les textes, les tableurs, les courriels, etc.]

L’arbre de décision a d’abord été formulé par le comité du programme et raffiné ensuite par Tom Strang, de l’ICC, et Bruce Walton, de BAC. Les choix les plus importants que doit faire une institution et les décisions clés qu’elle doit prendre au moment de penser à l’acquisition et à la préservation des documents électroniques ont mené à l’établissement de cinq catégories principales de séances : Évaluation des documents électroniques; Authenticité des documents électroniques; Élaboration d’une stratégie de préservation des documents électroniques et Connaissance des supports. Cette approche a été soulignée dans la demande de communications, et plus de 45 résumés ont été soumis. Grâce à la structure du programme en arbre de décision, le processus de sélection a été très précis. Le programme final a consisté en 29 communications représentant six pays : le Canada (14); les États-Unis (8); l’Australie (3); le Royaume-Uni (2); la France (1) et l’Allemagne (1). En plus du programme principal, quatre affiches qui cadraient avec la structure du programme furent également présentées (une du Canada, deux des États-Unis et une de la France).<sup>1</sup>

Un autre objectif du programme consistait à présenter une grande variété de petites à moyennes institutions possédant non seulement des archives et des

---

1. Une affiche présentée au symposium portait sur l’arbre de décision qui apparaît à la page 2 de l’introduction du présent ouvrage. Par conséquent, seules trois affiches font partie de la section sur les affiches.

of the electronic records typically found in archives and libraries. Thus, the speakers list included several individuals from the museum and gallery communities.

Finally, it was important that the program include some case histories that highlighted what various institutions have actually done or are doing to preserve electronic records. While discussing preservation strategies is important, seeing strategies actually being implemented and working in the real world provides useful information and models to follow. Among the case histories presented were the preservation of audio language recordings from Aboriginal elders in the Northwest Territories in Canada, the preservation of American poet Robert Creeley's computer files, and the implementation of the Victorian Electronic Records Strategy (VERS) within the Government of the State of Victoria, Australia. The case histories were arguably the most useful aspect of a program that included something for everyone.

As a whole, the program provided delegates with a sense of the broader issues involved in collecting and preserving electronic records, as well as knowledge about the challenges that other institutions are facing and how they are dealing with them.

We are pleased to present the papers from the symposium program in this book of postprints. However, because speakers were given the opportunity to revise their papers after the symposium and the final submissions were lightly edited, the text herein may differ slightly from the original presentations.

*Jane Down, CCI*  
*Richard Green, LAC*  
*Joe Iraci, CCI*  
*Mary Murphy, LAC*  
*Patricia Young, CHIN*

*Note: The papers in this book are published in the language in which they were presented, but all include an abstract in both English and French.*

bibliothèques, mais aussi des institutions culturelles telles que des musées qui doivent conserver des documents électroniques. Les musées, par exemple, incluent souvent de l'art vidéo dans leurs collections, mais les besoins de consultation de ces documents électroniques sont bien différents de ceux qu'on retrouve habituellement dans les archives et les bibliothèques. Ainsi, la liste des conférenciers comprenait plusieurs personnes du monde muséal.

Enfin, il était important de retrouver dans le programme des histoires de cas afin d'illustrer ce que diverses institutions ont fait ou font concrètement pour préserver leurs documents électroniques. La discussion des stratégies de préservation est certes importante, mais le fait de connaître des stratégies qui sont mises en œuvre et qui fonctionnent dans une situation réelle permet d'obtenir de l'information utile et d'avoir des modèles à suivre. Parmi les cas décrits, il y a eu la préservation des enregistrements sonores des aînés autochtones des Territoires du Nord-Ouest, au Canada, la préservation des fichiers informatiques du poète américain Robert Creeley et la mise en application de la stratégie VERS pour les documents électroniques de l'État de Victoria, en Australie. On peut soutenir que les cas présentés formaient l'aspect le plus utile d'un programme répondant aux attentes d'un grand nombre de participants.

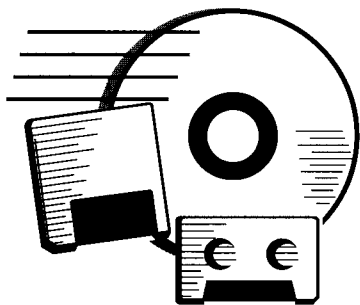
En général, les délégués sont repartis avec une vue d'ensemble des grandes questions liées à la collection et à la préservation des documents électroniques de même qu'une connaissance des problèmes auxquels se heurtent d'autres institutions ainsi que de leur manière de les résoudre.

Nous sommes fiers de vous présenter les communications du symposium dans cet ouvrage. Cependant, étant donné que les conférenciers ont eu l'occasion de modifier leurs textes après le symposium et qu'ils ont été légèrement révisés, il est possible que les textes soient quelque peu différents de ceux présentés durant le symposium.

*Jane Down, ICC*  
*Richard Green, BAC*  
*Joe Iraci, ICC*  
*Mary Murphy, BAC*  
*Patricia Young, RCIP*

*Remarque : les communications sont publiées dans la langue utilisée lors de la présentation, mais toutes sont accompagnées d'un résumé en français et en anglais.*





# DIGITAL PRESERVATION STRATEGIES: THE INITIAL OUTCOMES OF THE ERPANET CASE STUDIES<sup>1</sup>

SEAMUS ROSS

Director, Electronic Resource Preservation and Access Network (ERPANET)  
Humanities Advanced Technology and Information Institute (HATII)  
University of Glasgow  
Glasgow, Scotland  
United Kingdom  
E-mail: s.ross:@hatii.arts.gla.ac.uk

MONICA GREENAN

British Content Editor ERPANET  
(May 2002–April 2003)  
HATII  
University of Glasgow  
Glasgow, Scotland  
United Kingdom  
E-mail: british.editor@erpanet.org

PETER MCKINNEY

Coordinator ERPANET  
HATII  
University of Glasgow  
Glasgow, Scotland  
United Kingdom  
E-mail: p.mckinney@hatii.arts.gla.ac.uk

## Abstract

The Electronic Resource Preservation and Access Network (ERPANET) is a key European Commission Fifth Framework digital preservation activity. Among its other activities, it has recently conducted a series of case studies to improve its knowledge of digital preservation practices within Europe's public institutions and private-sector companies. Information for the case studies was collected between August and November 2002 by interviewing staff from 20 companies spread across the broadcasting, pharmaceutical, publishing, and telecommunications sectors. This paper describes the methods used to conduct this research and presents the results; it also defines areas where further work (e.g. research, guidelines, improved practices) is urgently needed.

## Introduction

Records managers, archivists, and librarians all recognize the need for urgent action to ensure the survival of digital information. In order to understand how European organizations are approaching digital preservation, ERPANET<sup>2</sup> has been conducting case studies to investigate awareness, strategies, and technologies. More than 500 organizations, institutions, and public bodies will eventually contribute to this research — even if for some it may only be by taking the decision not to take part.<sup>3</sup> The results should make a substantial

contribution to current 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 sectors included, but also through the depth at which the issues will be explored.

ERPANET is deliberately and systematically approaching disparate organizations from industry and business to facilitate discussion in areas that have traditionally been unconnected. With these case studies, ERPANET will be able to broaden the scope and understanding of digital preservation through research and discussion. The case studies will be published online on the ERPANET Web site<sup>4</sup> to help 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, for example: what can the publishing sector learn from the aeronautical sector? Eventually we aim to use this comparative data to produce cross-sectoral overviews. This cross-sectoral fertilization is a main focus of ERPANET as laid out in its Digital Preservation Charter.<sup>5</sup>

## Aims

The principal aims of the case studies are to:

- build a picture of methods and match them 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, through the identification of tools that are needed and the accumulation of information about good practices
- create material for training seminars and workshops
- develop contacts

## Methodology

Potential sectors on which to base the case studies are selected to represent a wide scope of information production and digital preservation activity. Each sector presents a unique perspective on digital preservation, as organizational and sectoral requirements, awareness of digital preservation, resources available, and the nature of the digital object created place unique and specific demands on organizations. Initial desk-based sectoral analysis provides ERPANET researchers with essential background knowledge. The primary research is then conducted 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 of the Pittsburgh Project<sup>6</sup> and InterPARES I<sup>7</sup> for target selection and data collection were given special attention. But after careful consideration, the Pittsburgh approach was considered too narrowly focused and provided insufficient breadth to enable sectoral comparisons. And the InterPARES I data collection methodology proved too detailed and lengthy (which could become an obstacle both at the point of data collection and during analysis) and focused too closely on record-keeping systems (we wished to take a wider view).

The ERPANET interview instrument that was eventually developed (see Appendix<sup>8</sup>) takes into account both the strengths and weaknesses of earlier projects, and targets a range of strategic points in the organizations under examination. It explores three main areas of enquiry within an organization: awareness of digital preservation and the issues surrounding it; digital preservation strategies (both in planning and in practice); and future preservation requirements. Within these three themes, distinct layers of questions uncover the state of the entire digital preservation process within the institution.

Drawing on the experience that the partners of ERPANET have in this method of research, another feature is also included, i.e. wherever possible three categories of employee within each organization are interviewed: an Information Systems or Technology Manager, a Business Manager, and an Archivist/Records Manager.<sup>9</sup> This provides two members of staff with knowledge of the organization's digital preservation activities, and a manager to offer an overview of business and organizational issues. When completed, this interview instrument provides information on the extent of knowledge and practice in organizations, who has responsibility and problem ownership, and where the drive towards digital preservation is initiated within organizations.

The main areas of investigation include:

- perception and awareness of risk associated with information loss
- level of understanding of how digital preservation affects the organization
- actions that have been taken to prevent data loss
- processes for monitoring preservation activity
- mechanisms for determining future requirements

Within each section, the questions are designed to bring organizational perceptions and practices into focus. Questions are aimed at understanding impressions held on digital preservation and the impact that it has on the respective organization, exploring the awareness in the sector of the issues and the importance that digital preservation is accorded, and determining how digital preservation affects organizational thinking. For instance, participants are asked to describe (in their views) what the main problems associated with digital preservation are and what long-term value information actually has in their sector. Through this the reasons for preserving information as well as the risks associated with not preserving it become clear.

The core of the questionnaire focuses on the actions taken at corporate 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 allow participants to explore the future commitment from their organization and sector to digital preservation activities and, where possible, to relate their existing or planned activities to those being conducted in other organizations with which they might be familiar.

The first round of the case studies focused on companies drawn from the pharmaceutical,

broadcasting, publishing, and telecommunications sectors. These sectors were chosen to guarantee a diversity of organizational type, activity, regulatory framework, and culture. The aim was also to include a range of institutional types, sizes, and locations in the case studies, although this was not always possible (e.g. in the telecommunications sector).

This survey process was not without its problems. Even when organizations were identified and interviews time-tabled, targets often withdrew just before we began the interview process. Some withdrew after seeing the data collection instrument and realizing the time/effort involved, and others (we suspect) because they realized that the expertise was not available within their organization to answer the questions. In the latter instance, withdrawal might indicate that the organization was not maintaining their digital records effectively. The perception of risks that might arise through contributing to the study worried some organizations, particularly those from sectors where competitive advantage is imperative, or liability and litigation issues especially worrying.<sup>10</sup> Telecommunications proved to be one of the most difficult. Non-disclosure agreements that stipulated that we would neither name an organization nor disclose any information that would enable readers to identify them without their explicit approval were used to reduce risks associated with contributing to this study. In some cases the risk was still deemed too great and organizations withdrew. Several organizations withdrew on seeing the final report because its conclusions were unsettling (one particularly worrying instance from a later set of case studies was the withdrawal of a major national nuclear energy producer on sight of their report).

## **Overview of findings from the first round of studies**

This first round of case studies was undertaken between August and November 2002. Nearly 50 staff<sup>11</sup> from 20 organizations in the pharmaceutical,<sup>12</sup> broadcasting,<sup>13</sup> publishing,<sup>14</sup> and telecommunications<sup>15</sup> sectors were interviewed. Findings from the case studies showed wide variation in current strategies and practices. The organizations surveyed had a huge amount of digital information to handle, and the types of digital objects themselves varied across the sectors. A wide variety of reasons and justifications for maintaining this information existed. Diversity in preservation activity reflected the different levels of understanding of the issues and a lack of consensus on the best methods to meet preservation challenges.

## **Perception and awareness**

An assessment of the levels of awareness of both digital preservation issues and the value of information was included to gauge the depth of understanding and the relevance of long-term access to digital materials to the individual organizations. Uncovering what preservation issues were most prominent for both the organizations and the sectors provided a mechanism to assess the relevance of digital preservation for them.

The two sectors with the deepest awareness were the pharmaceutical and broadcasting sectors, with the broadcasting sector in particular displaying an acute understanding of the challenges of digital preservation. This is a sector that actively pursues solutions collaboratively. This awareness reflects their widespread appreciation of the re-use value of broadcasting materials. The pharmaceutical sector has had digital preservation awareness imposed by external regulators, who have created a complex and challenging structure of compliance and an economic advantage for the filing of drug approval requests in digital form. The driving force behind the development of solutions by the pharmaceutical sector is the legal requirement to preserve information relating to many of its products, although the fiercely competitive nature of the individual organizations within the sector has resulted in these solutions being developed in isolation. Similarly, the preservation activities of the telecommunications sector are driven by legal requirement rather than the desire to generate historical records. Thus data on call records, contracts, and interconnect data are kept for only 10 years. The publishing sector has become increasingly aware of the problem in the last few years, especially with the growth in electronic journal publication. Publishers are still attempting to establish who is responsible for the preservation — the publishers or the libraries (which have, in the print environment, traditionally played a central role in this activity).

It is not really surprising that the values attributed to information by organizations and sectors varied. Where most of the organizations fell down, however, was in their failure to transform their digital material into assets.<sup>16</sup> Respondents typically classified information as either having a re-use value or a cultural or historical value. However, only the most established publishing organizations and some of the broadcasting organizations were concerned about preserving information for its cultural or historical value; a substantial amount of what is produced in these sectors is considered a cultural asset. In the pharmaceutical sector, there were two clear opinions on the re-use value of information: those who

believed that scientific research information can be re-used (the expense of creating research data and the vast amounts created were the main reasons for this attitude); and those who felt that if re-investigation had to be carried out in the future, technological advances would make the existing data obsolete. Within the publishing sector, some of the older publishers were concerned about the historical value of information, but for most the re-use value was the most prominent issue: access to publications was essential to enable their republication and to create multiple ways to exploit the products. However, within the publishing sector, individuals often attributed value to digital materials in the absence of a clear organizational position.

### **Cooperation**

By including a focus on knowledge sharing, the case studies aimed to discover current and future levels of cooperation.

Results showed that in most cases collaborative effort to tackle the problems of digital preservation was rare. However, within the broadcasting sector, a lot of work has been undertaken to develop standards and best practices for the management and preservation of broadcasting material. International organizations such as the European Broadcasting Union (EBU)<sup>17</sup> and the European Commission have funded research projects such as the Preservation Technology for European Broadcast Archives (PRESTO)<sup>18</sup> which have fostered the development of strategies for the preservation of film, video, and audio material. Efforts have also been channeled into the exchange of information, cooperation on shared projects, and professional training. There is some discussion among the pharmaceutical sector but a specific focus on digital preservation is rare — and when discussions do take place there is little in the way of detail due to the intense competition inherent in this sector. Pharmaceuticals expend large amounts of resources on tackling preservation problems by themselves without much concern for the techniques and strategies developed by competitors or companies from other sectors. Common initiatives are also unusual in the publishing sector, but publishers are at least aware of what is going on in other organizations. Many report that they anticipate collaborative work with libraries as the next likely development; the collaboration between the National Library of the Netherlands (Koninklijke Bibliotheek) and Reed Elsevier is an example of the potential here.<sup>19</sup>

### **Policy, strategy, and standards development**

In identifying current practice within organizations, the case study attempted to uncover the variety of

policies, strategies, and standards that have been developed, adopted, and implemented.

All of those interviewed were aware of the need to identify and implement solutions, and develop them.<sup>20</sup> For those that had developed specific strategies and policies, these were often not implemented across all departments and satellites.<sup>21</sup> Broadcasters indicated that they exchange information on costing policies, strategies to develop technical solutions, standards, and implementation guidelines. This effort is reflected internally with directives, standard procedures, and preservation programs dealing with preservation requirements, recovery, standard format, and metadata systems. For other sectors, policies are developed internally with additional support from external service providers. Most pharmaceutical organizations reported that they plan to develop preservation policies as part of their projects articulating system requirements and specifications to software suppliers and system developers. Among our cohort were some who reported that their organizations do not plan to adhere to standards and guidelines established by other communities, although they do intend to review other policies and standards and build on them. Publishers have few strategies in place, and this may reflect a lack of systematic approaches to preservation. Most indicated that a good strategy was to keep everything; this, in their view, at least ensured that the material would be there in the future. But the approach begs questions about documentation, formats, and curation. There is a need, therefore, for a general raising of awareness of appraisal strategies.

### **Selection of digital information for preservation**

The results indicate that most staff members have a solid theoretical awareness of the issues and the need to determine retention. This does not, however, mean that this knowledge is being put into practice. Within the publishing sector, if there is a need to keep information, this is done within an organization on an individual or departmental basis. Often it is those with a stake in the information and its re-use value who act to secure it rather than the organization itself. This ad hoc practice has operated successfully in the past; hence there is little impetus for the adoption of more formal controls. Pharmaceutical organizations preserve all classes of information that regulating authorities stipulate must be retained. In this case, selection criteria from the paper world have been transferred to the digital. As a result of falling storage costs, many organizations do not consider it worth the effort to filter digital information but instead retain all data. Selection criteria in the broadcasting sector take into account archival, technological, and

practical considerations, with particular emphasis on re-use value. Internal directives that also set classification and management procedures of retention schedules regulate selection criteria. Only a few organizations indicated that any sort of selection policy had been developed in consultation with internal departments and organizational units with a stake in the preservation of digital objects.

### **Preservation and storage**

Within the pharmaceutical environment, the size and proliferation of formats are the main obstacles to the preservation of objects. Overall the interviewees indicated that their organizations show little impetus to develop solutions to ensure digital longevity: the majority believe that solutions should be developed by external, specialized suppliers in order to allow organizations to concentrate on their actual field of operations and area of expertise. A pragmatic approach is the one most generally taken, with the migration to new data formats undertaken when the need arises, although few organizations seemed to be aware of the complexities associated with migration. To be safe, some organizations retain the digital object in its original format alongside the migrated version. In regulated sectors, such as pharmaceuticals, authenticity, integrity, and confidentiality of the records must be guaranteed. Validating these elements at the level of the individual digital object would be prohibitively expensive. The optimum point of validation is system level, with support of transaction audit trails. For the digital objects themselves, pdf (portable data format) has emerged as the prevalent preservation format, with several organizations restricting the use of its special features to enhance its effectiveness. A technique for the long-term preservation of more complicated or specialized data formats remains to be adequately addressed. The effort involved in convincing suppliers of scientific equipment to use standard formats was seen as too great due the time that would need to be spent on developing such agreements. Of the four sectors, it is perhaps the pharmaceutical sector that has sufficient power to promote a change in attitude on the part of suppliers. Although publishing and broadcasting together exercise greater financial muscle (as measured by annual turnover) than the pharmaceutical sector, issues of liability and risk do not play as significant a role for them.

Responsibility for preservation activity in the publishing organizations falls on the information technology departments. Results indicated that pdf is a very popular format for distribution and preservation; other formats include tiff (Tagged Image File Format), xml (Extensible Markup

Language), and sgml (Standardized General Markup Language). Most organizations noted that they are waiting for an industry format to emerge. Preservation planning is made either as an ad hoc decision or because of a vendor relationship. Some organizations still consider retention of a print or microform copy as the most reliable preservation option currently (2002) available.

The participating telecommunications organizations reported that their sector integrates preservation activities into their business software, and migrations are conducted with every update to newer versions of the software. As 10 years is the maximum retention period, this approach is considered sufficient. But no standard metadata sets were applied, and there was little appreciation of their value. One organization interviewed did transfer expired files onto CD-R for preservation, but only as a matter of historical interest — and no resources were expended on monitoring the material or maintaining its accessibility after the initial transfer.

For broadcasters, results indicated that preservation activities are either carried out entirely in-house or partially outsourced as a way of reducing costs, and training of personnel is considered fundamental to the process. Most of the material produced for programming is converted to digital format, if not already in that state. Metadata production and collection is a central part of the access and preservation system, and a substantial investment of human and financial resources is deployed in this area. The formats currently used by this sector are considered reliable and migration is not seen as an immediate issue. Out of all the sectors explored, broadcasting sees the greatest need for the long-term preservation of its materials; one participant suggested that an international repository that would take responsibility for digital materials should be set up.

### **Digital preservation costs**

The cost of digital preservation concerned all organizations interviewed. Some see the financing of digital preservation activities as corporate investment, whereas others pass responsibility to research and development divisions. This dichotomy exists because of the division of responsibilities. In a few organizations the responsibility for preservation is shared by all parts of the organization, and hence the costs are spread across the stakeholders. In these organizations, viable and manageable cost mechanisms usually exist.

The broadcasting sector expends substantial resources on digital preservation, and is now

trying to streamline activities in order to rationalize its spending. They will concentrate on continuing digitization efforts, providing more access (e-delivery) and developing mass storage solutions. Where there is a specific budget for digital preservation, it commonly amounts to 1% of the organization's total budget. Allocations are identified through an accurate cost-and-benefit analysis based on previous experience and the results of various initiatives. Occasionally broadcasters are able to secure additional external funding because their digital objects are considered to be cultural assets. The availability of European Commission funding has been a major factor in fostering cooperation among broadcasters.<sup>22</sup> Interviewees from the publishing sector reported that their companies would like to spend more on digital preservation, but recognize that a better understanding of the costs involved has to be established before any investment will be possible. In order to increase spending, the budget holders must be convinced by a solid business case. Pharmaceutical organizations were concerned by the risks involved in spending on digital preservation, but felt that the current level of expenditure was broadly right. A few wished to spend more on preserving the vast quantities of information resulting from research and development activity for its research potential.

#### **Monitoring of actions**

With strategies and techniques in place, auditing procedures offer crucial mechanisms to measure compliance over time. Compliance with legal requirements and, where in existence, organizational directives, were common ways of monitoring activity. In more than 50% of the target organizations there was little or no monitoring. This reflected the lack of formalized preservation processes in the first place. Aside from internal monitoring, pharmaceutical organizations are subject to external inspections by regulatory agencies to audit their compliance. It is compliance and risk that dictate the processes that the pharmaceutical sector has undertaken, and is planning to undertake. Less regulated sectors have not been as quick to address the preservation problems.

#### **Future requirements**

Most participating institutions indicated that they believe their current state of activity is sufficient in the short term, and they are confident that their organization will be able to deal with any problems that arise in a manner such that information will not be lost and the organization will not suffer any detrimental effects, including financial loss or increased risk. To ensure that this will indeed be true, the need to preserve an increasingly large quantity of records and information should be linked to a

business case to improve and expand access to it. Some sectors reported that solutions in such areas as mass storage and migration would be best procured from external commercial suppliers. Both the publishing and the telecommunications sectors were keen to see increased development of clear and implementable organizational strategy and policy, with commitment and support from the senior levels of management. The publishing sector in particular indicated their desire to increase the number and competencies of personnel dedicated to digital preservation within organizations. Pharmaceutical organizations would like to see the disparate preservation activities and storage brought together in an organization-wide solution.

## **Conclusions**

This paper has drawn attention to the diverse nature of digital preservation strategies and practices. It has also examined the differences in requirements between sectors, but despite these differences all institutions would still benefit from the intra- and cross-sectoral sharing of knowledge. Most of the organizations, with the exception of the broadcasting sector, are hesitant in their digital preservation activities. They are waiting for external developments that they can adopt, or off-the-shelf solutions that they can implement. Common responses indicate that, due to the fragmentary nature of most of these organizations, little is undertaken on an organizational basis; instead, responsibility for preservation is shifted down either to department level or, in many cases, the individual.

Hesitancy arises from a lack of awareness in some sectors. Not only in terms of the enormity of the problem, but also in terms of solutions that have been developed in other areas. Indeed, one of the respondents indicated that when first approached by ERPANET about participating, the initial contact e-mail passed through the hands of ten people before someone was found in the organization who understood the problem well enough to get involved.

The broadcasting example indicates that strong inter-institutional cooperation leads to increased confidence and, in turn, increased activity and development of common knowledge and understanding. Furthermore, it is widely regarded that cooperation will lead to the development of more robust and financially viable tools and solutions. The pharmaceutical sector might learn from the success of the broadcasting sector, but whether their inherent risks in sharing information would outweigh any benefit would need to be assessed. The successes

in the telecommunications sector, where actions and responsibility are spread around the organization, offer a streamlined model that others would benefit from implementing.

The case studies to date have outlined the knowledge, requirements, skills, and practices in digital preservation. However, communication and a common platform for information sharing are also key to the success of individual organizations' strategies. Recommendations from this round of studies are that intra-sectoral collaboration take place and that cross-sectoral discussion be encouraged. There are definitely areas where sharing can occur — strategies, costing models, and general awareness. From these case studies ERPANET has identified five areas where it will promote further action:

- creation of off-the-shelf policy statements
- development of business cases and strategies
- drafting of clear guidance on technologies and their preservation implications
- constructions of improved models (reference, costs, standards, functional requirements)
- production of guidance on creating, managing, and auditing digital repositories

It is with this in mind that ERPANET began in mid 2003 to develop briefing papers laying out guidelines that will enable organizations to measure their "preservation effectiveness" and improve their digital preservation practices. We are also beginning to engage in discussion with software vendors aimed at convincing them of the importance of incorporating preservation functionality into their products. Through these activities we hope to promote the development of cross-sectoral strategies for digital preservation.

## Endnotes

1. The authors would like to thank their ERPANET colleagues Andreas Aschenbrenner (Nationaal Archief Nederland, Den Haag), Georg Büchler (Schweizerisches Bundesarchiv), and Prisca Giordani and Vincenzo Di Meo (ISTBAL, Università di Urbino) for helping to conduct the case studies.
2. ERPANET is a European Commission funded activity (IST-2001-32706). The Directors of ERPANET are Niklaus Bütikofer (Schweizerisches Bundesarchiv), Maria Guercio (ISTBAL, Università di Urbino), Hans Hofman (Nationaal Archief Nederland, Den Haag), and Seamus Ross (HATII, University of Glasgow). See the ERPANET Web site ([www.erpanet.org](http://www.erpanet.org)) for more details and available products.
3. Since the original paper was submitted, ERPANET has decided to conduct a broader survey of leading companies in the 15 EU Member States and the 10 countries that are about to join the EU. See the ERPANET Web site ([www.erpanet.org](http://www.erpanet.org)) for more information.
4. Case studies are made available on the ERPANET Web site ([www.erpanet.org](http://www.erpanet.org)) as soon as they are completed.
5. The Charter is ERPANET's statement on the principles of digital preservation. It has been drafted to achieve a concerted and coordinated effort in the area of digital preservation by all organizations and individuals that have an interest in and share these concerns. It can be viewed on the ERPANET Web site ([www.erpanet.org/www/content/documents/Digitalpreservationcharter4\\_1.pdf](http://www.erpanet.org/www/content/documents/Digitalpreservationcharter4_1.pdf)).
6. The Web site of the Pittsburgh Project was lost due to a "technical glitch." However, interested individuals are encouraged to go to the Internet Archive site ([www.archive.org/](http://www.archive.org/)) and use the "Wayback Machine" to find the Pittsburgh Project ([www.sis.pitt.edu/~nhprc](http://www.sis.pitt.edu/~nhprc)). Our evaluation was carried out using a hard copy of the materials produced by the Pittsburgh Project which was provided to Seamus Ross in 1995 for his review *Commentary on the Pittsburgh University Recordkeeping Requirements Project: A Progress Report (Delivery Draft)*, Society of American Archivists, 59th Annual Meeting, Washington, DC (Thursday 31 August 1995, Session 20); this review is available at ERPAePRINTS ([eprints.erpanet.org/](http://eprints.erpanet.org/)).
7. [www.interpares.org](http://www.interpares.org)
8. We include the interview instrument to encourage comment and in the hope that others conducting similar research might adopt it (or a variant of it) as a way of increasing the comparability between different studies.
9. See the individual case study reports ([www.erpanet.org](http://www.erpanet.org)) as this was not always possible.
10. It has been suggested that ERPANET should present the results of the studies as anonymous reports (e.g. An International French Bank, a European Nuclear Generator), because the participation rate might increase if companies knew they would not be specifically named. The Directors considered this approach, but concluded that case studies with identifiable targets have authenticity and "realism" that anonymous reports lack.

11. The original target of three staff per organization was not always achieved; see the individual case study reports ([www.erpanet.org](http://www.erpanet.org)) for an explanation.
12. Astra Zeneca PLC ([www.astrazeneca.com/](http://www.astrazeneca.com/)); Aventis Pharma Germany GmbH ([www.pharma.aventis.de](http://www.pharma.aventis.de)); Bayer AG ([www.bayer-pharmaceuticals.com](http://www.bayer-pharmaceuticals.com)); Böhringer Ingelheim ([www.boehringer-ingelheim.com](http://www.boehringer-ingelheim.com)); GlaxoSmithKline plc (GSK) ([www.gsk.com](http://www.gsk.com)); Organon NV ([www.organon.nl/](http://www.organon.nl/)); Pfizer ([www.pfizer.com](http://www.pfizer.com)). Fourteen other pharmaceuticals were contacted and, although a number expressed initial interest in taking part, they all eventually decided not to contribute to the study.
13. RAI (Radiotelevisione Italiana) ([www.rai.it](http://www.rai.it)); BBC (British Broadcasting Corporation) ([www.bbc.co.uk](http://www.bbc.co.uk)); ERT (Elliniki Radiophonia Tileorassi SA) ([www.ert.gr](http://www.ert.gr)); SF DRS (Schweizer Fernsehen der Deutschen und Rätoromanischen Schweiz) ([www.sfdrs.ch](http://www.sfdrs.ch)).
14. Gruppo Mondadori ([www.monadori.com](http://www.monadori.com)); Hachette Filipacchi Médias ([www.hachette-filipacchi.com](http://www.hachette-filipacchi.com)); HMSO (Her Majesty's Stationary Office) ([www.hmso.gov.uk](http://www.hmso.gov.uk)); Kluwer Academic Publishing ([www.wkap.nl/](http://www.wkap.nl/)); Office for Official Publications of the European Communities (Publications Office) ([eur-op.eu.int/general/en/index\\_en.htm](http://eur-op.eu.int/general/en/index_en.htm)); Oxford University Press ([www.oup.co.uk/](http://www.oup.co.uk/)); Random House ([www.randomhouse.com/home.pperl](http://www.randomhouse.com/home.pperl)).
15. Swisscom ([www.swisscom.com](http://www.swisscom.com)); Orange ([www.orange.com](http://www.orange.com)). Numerous other telecommunications companies across Europe were approached, but most refused to take part or dropped out. When preparing this paper we considered dropping the telecommunications sector because it included too few companies, had a poor geographic spread, and did not include enough variation in organizational size. However, because the results from the telecommunications sector were consistent with the results from other sectors, we decided that it would be beneficial to include them.
16. For more on this see: Moon, M. "How do Cultural Artefacts become Digital Assets?" *Digital Asset Management Systems for the Cultural and Scientific Heritage Sector. DigiCULT Thematic Issue 2* (2002).
17. [www.ebu.ch/](http://www.ebu.ch/)
18. [presto.joanneum.ac.at/index.asp](http://presto.joanneum.ac.at/index.asp)
19. [www.kb.nl/kb/resources/frameset\\_kb.html?/kb/pr/pers/pers2002/elsevier-en.html](http://www.kb.nl/kb/resources/frameset_kb.html?/kb/pr/pers/pers2002/elsevier-en.html)
20. The ERPANET Seminar on Policies and Procedures held in Fontainebleau in January 2003 added further weight to the findings of these case studies. There was a demand for guidance in the development of policies, and even those organizations that had them in place acknowledged that they were not always implemented. For more information see the ERPANET Web site ([www.erpanet.org/www/products/Paris/paris.htm](http://www.erpanet.org/www/products/Paris/paris.htm)).
21. Even where policies are in place they are often unrealistic and, as a result, unimplementable.
22. Projects such as COLLATE (Collaboratory for Annotation, Indexing, and Retrieval of Digitized Historical Archive Material) ([www.collate.de](http://www.collate.de)), AMICITIA (Asset Management Integration of Cultural heritage In The Interexchange between Archives) ([www.amicitia-project.de/index.shtml](http://www.amicitia-project.de/index.shtml)), PRESTO (Preservation Technology) ([presto.joanneum.ac.at/index.asp](http://presto.joanneum.ac.at/index.asp)), and FIRST (Film Restoration and Conservation Strategies) ([www.film-first.org/emf/first.asp?cod\\_seccao=208](http://www.film-first.org/emf/first.asp?cod_seccao=208)) were made possible under the IST (Information Society Technology) Fifth Framework Programme.



## Appendix: ERPANET case study data collection instrument

Administrative Section  
Interview Details  
Organisation Details  
Disclosure/Privacy Information  
Tracking of Activities

### **Perception and Awareness of Digital Preservation**

We would like to begin by asking you a few questions about your general impressions of digital preservation, and the impact that it has on the \_\_\_\_\_ sector. We will use the term 'digital information' throughout to refer to all forms of digital data, records and information.

1. Is there a general awareness in the \_\_\_\_\_ sector that the long-term preservation (more than five years) of digital information is an important issue?
2. To what extent does the sector recognise the importance of preserving digital information in the long term?
3. What are the main problems associated with digital preservation in the \_\_\_\_\_ sector?
4. From what sources have you heard about the issues surrounding digital preservation?
5. What values does digital information have in the \_\_\_\_\_ sector beyond the original purposes for which it was created?

### **Understanding How Digital Preservation Affects Your Organisation**

We would like to focus on how some of these digital preservation issues affect your own organisation.

6. What type of information is digitally preserved in the short and the long term in your organisation?
7. What are the reasons that digital information is preserved in your organisation?
  - Legal requirements
  - Financial requirements
  - Business requirements (e.g. document important decisions and activities)
  - Historical value
  - Other (please specify)

8. What risks is your organisation under if digital information is not preserved in the long term?
  - Legal risks
  - Financial risks
  - Business risks
  - Historical value
  - Other (please specify)
9. Has the organisation conducted a risk analysis and/or business needs analysis with regard to the preservation of information? If yes, can you indicate the main results?

### **Actions Taken: Policies, Strategies, Standards and Practices Developed**

The questions in this section aim to explore some of the actions that the organisation has undertaken to deal with the preservation of electronic records. It will examine the above as well as selection, preservation, storage, and access activities.

#### Policies, Strategies, and Standards

10. Is there any collaborative effort across the \_\_\_\_\_ sector to tackle common digital preservation issues?
  - Conferences
  - Newsletters
  - Journals
  - Common institutions
  - Collaborative projects
  - Other (please specify)
11. Has your organisation attempted to find information external to the sector regarding preservation?  
If yes, please indicate the sources.
  - Government agencies
  - Higher education institutions
  - Archives
  - Libraries
  - Museums
  - IT specialists
  - Other (please specify)

Please specify the kind of information provided and how useful it proved to be.

---

---

---

---

---

12. Do you cooperate with other institutions in the research and development of policies, strategies, and standards? In what way?

Yes  
 No  
If yes, please specify.

13. How useful is this common effort in applying it to your organisation's own needs?

---

---

---

---

14. Do you have any specific organisational policies that relate to the preservation of information?

15. Who (and what) was/is involved in the creation of these policies?

- Management
- Employees
- Special task force in the organisation
- Results of internal analyses (e.g. risk analysis)
- External sources, models, advice
- Other (please specify)

23. Is your classification and retention schedule linked and implemented across the organisation?

24. Who is responsible for the maintenance and implementation of these schedules?

25. How do you ensure that selected information is complete, accurate and identifiable?

### **Preservation of Digital Information**

16. Do these policies apply across the entire organisation?

26. Does your organisation take care of its preservation activities itself, or are these outsourced?

17. How are these policies implemented?

- Outsourced
- In-house

18. Has your organisation developed preservation strategies, standards, and practices and implemented them?

- Yes
- No

If yes, please specify.

If outsourced, what reasons were behind this decision, and who carries out the preservation activities?

---

---

---

---

---

---

---

---

---

---

19. How were they introduced and implemented (e.g. by department, with training)?

27. Are there specific individuals in your organisation responsible for the preservation of digital information?

20. How, and under whose responsibility have these been established?

- External advice/sources/models
- Survey of information resources
- In-house solutions developed
- Other (please specify)

28. What positions do these people hold in the organisation, and what are their responsibilities and competencies?

29. What type of training or advice is available for them?

21. How often are your preservation policies and strategies updated and renewed?

30. Is your organisation aware of any external standards, best practices, and guidelines available on preservation?

- Yes
- No

If yes, please specify.

---

---

---

---

### **Selection of Digital Information for Preservation**

22. Do you have a selection policy, or classification and retention policy that determines what information in your organisation is to be preserved?

31. Are these specific to your sector?

32. Where did you learn about them? Please specify your sources.
33. Which of these standards, practices and guidelines do you use?
34. What technologies do you use for preservation? For the following list of current techniques, please specify which ones you use and for what kind of information.

Technique	Specify type/ technology used	Information preserved
Print to paper		
Scanning		
Save on disk		
Save on other media		
Emulation		
Migration		
Microfilm/microfiche		
Other		

35. On what grounds were these techniques chosen? Please specify your answers.
- External advice
  - Trials and evaluations
  - Recommendations
  - Intra-sectoral standards available
  - Other
- Please provide as much information as possible about why these decisions were taken.

---



---



---



---

36. What data formats do you use for preservation?
- Standard data formats
  - Other
- Please specify for both answers.

---



---



---



---

37. Do you convert the information to be preserved into other data formats for technical (or other) reasons?
38. What metadata do you use to describe both your digital information and the processes of storage and preservation? Does it follow any standards available (Dublin Core or other)? Can you provide a copy of the metadata set?

39. Is the collection and production of metadata automated?
40. Who is responsible for the transfer of information into long-term storage?
41. How often (if undertaken) is digital information migrated or refreshed?

### **Storage of Digital Information**

42. Do you have a particular storage area for digital information to be preserved?

Yes

No

If yes, how is this organised and equipped?

---



---



---



---



---

43. Do you keep redundant copies of the digital information to be preserved for safety (or other) reasons?

### **Access to Digital Information**

44. How is information protected from inadvertent or unauthorised access and manipulation?

45. Does your preservation solution allow direct access to the digital information stored (i.e. are they stored in an executable format)? If no, how is the access provided?

46. What access issues does your organisation face?
- a. Copyright
  - b. Privacy issues
  - c. Access security and privileges
  - d. Other (please specify)

47. How does your organisation intend to provide access to digital information into the future?

### **Digital Preservation Costs**

48. Did your organisation attempt to undertake a cost benefit analysis concerning its investments in preservation?

49. Has this analysis been assessed in light of your actual preservation activities? Did it prove to be accurate?

50. To which section of the budget are the economic resources for your preservation programme allocated?

51. What percentage of the organisation's budget is spent on preservation? Can you compare that to some other area of the organisation's activity?

52. Is the organisation attempting to address amortisation issues in the preservation budget?

53. Are there available sources of funding within the \_\_\_\_\_ sector allocated for digital preservation issues?

Yes

No

If yes, please specify.

---

---

---

---

---

---

54. Are you satisfied with these cross-sector services?

55. If no, what would you like to see available (i.e. what would you think could best be solved in common in your sector)? Would you be willing to engage financially in such information?

56. Are there other external sources available for digital preservation activities (e.g. government grants, cross-sector funds)?

Yes

No

If yes, please specify.

---

---

---

---

---

---

## Monitoring of Actions

After having identified what has been undertaken in your organisation with regard to preservation activities, we would like to find out about how these efforts have been monitored.

57. Is the preservation process audited on a regular basis?

58. Is compliance to policies, standards, and strategies audited on a regular basis?

59. Is compliance to other requirements (legal, business etc.) audited on a regular basis?

60. How often are checks made to the preserved material (e.g. for signs of deterioration)?

61. Please specify the criteria used for these audits.

62. Who performs these audits (e.g. internal/ external)?

## Future Requirements

We would like to ask about the areas in which there is a need for additional attention in your organisation and the sector as a whole.

63. How long do you predict that your current preservation policies, strategies, and solutions will meet your organisation's preservation needs?

64. Is the amount of money allocated for preservation going to change in the future? Will it need to be changed?

65. If more funds were available, what could/would they be used for?

66. What conclusions has your organisation come to about its preservation efforts? Are these satisfactory?

67. What preservation efforts are remaining to be addressed within your organisation?

Further data to be preserved

Revision and adjustment of preservation policies and strategies

Additional resources dedicated to preservation

Technological solutions

Other (please specify)

68. Would you like to see more cross-sectoral or intra-sectoral activity with regard to preservation?

69. Are there any other areas in which you would like to have more information made available on digital information? Where do you expect this information to come from?

Thank you very much for your valuable contribution.

## **Résumé**

*L'Electronic Resource Preservation and Access Network (ERPANET) est une activité de préservation essentielle du 5<sup>e</sup> Programme-cadre de la Commission européenne. Au nombre de ses autres initiatives, le Réseau a mené récemment une série d'études de cas visant à améliorer sa connaissance des pratiques de préservation de documents numérisés dans les institutions publiques et le secteur privé en Europe. Les données qui ont servi à l'étude ont été recueillies entre les mois d'août et novembre 2002 dans le cadre d'entrevues auprès du personnel de 20 entreprises dans les secteurs de la pharmacie, de la radiodiffusion, de l'édition et des télécommunications. Cette communication décrit les méthodes employées pour effectuer cette recherche et en présente les résultats; elle précise également les domaines (recherche, lignes directrices, pratiques améliorées) où des efforts supplémentaires doivent être faits d'urgence.*